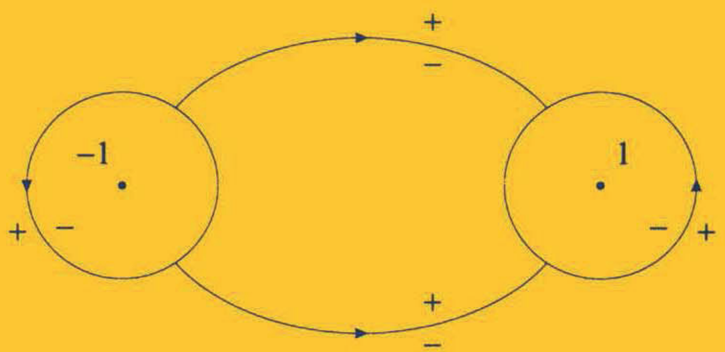
Lecture Notes in Mathematics

Erik Koelink
Walter Van Assche (Eds.)

Orthogonal Polynomials and Special Functions

1817

Leuven 2002



Springer

Lecture Notes in Mathematics 1817

Springer
Berlin
Heidelberg
New York
Hong Kong
London
Milan
Paris
Tokyo

Erik Koelink
Walter Van Assche (Eds.)

Orthogonal Polynomials and Special Functions

Leuven 2002

Springer

Editors

Erik Koelink

Afd. Toegepaste Wiskundige Analyse
EWI-TWA, Technische Universiteit Delft
Postbus 5031
2600 GA Delft, Netherlands

e-mail: h.t.koelink@math.tudelft.nl
http://fa.its.tudelft.nl/˜koelink

Walter Van Assche

Departement Wiskunde
Katholieke Universiteit Leuven
Celestijnenlaan 200B
3001 Leuven, Belgium

e-mail: walter@wis.kuleuven.ac.be
http://www.wis.kuleuven.ac.be/analyse/walter

Cataloging-in-Publication Data applied for

Bibliographic information published by Die Deutsche Bibliothek

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at http://dnb.ddb.de

Mathematics Subject Classification (2000): 33-01, 33F10, 68W30, 33C80, 22E47, 33C52, 05A15, 34E05, 34M40, 42C05, 30E25, 41A60

ISSN 0075-8434
ISBN 3-540-40375-2 Springer-Verlag Berlin Heidelberg New York

Springer-Verlag Berlin Heidelberg New York a member of BertelsmannSpringer Science + Business Media GmbH

http://www.springer.de

Typesetting: Camera-ready TEX output by the author

SPIN: 10933602 41/3142/ du - 543210 - Printed on acid-free paper

Preface

Orthogonal Polynomials and Special Functions (OPSF) is a very old branch of mathematics having a very rich history. Many famous mathematicians have contributed to the subject: Euler's work on the gamma function, Gauss's and Riemann's work on the hypergeometric functions and the hypergeometric differential equation, Abel's and Jacobi's work on elliptic functions, and so on. Usually the special functions have been introduced to solve a specific problem, and many of them occurred in solving the differential equations describing a physical problem, e.g., the astronomer Bessel introduced the functions named after him in his work on Kepler's problem of three bodies moving under mutual gravitation.

So the subject OPSF is very classical and there have been very interesting developments through the centuries, and there have been numerous applications to various branches of mathematics, e.g. combinatorics, representation theory, number theory, and applications to physics and astronomy, such as the afore-mentioned classical physical problems, but also integrable systems, optics, quantum chemistry, etcetera. So OPSF is well-established, and very much driven by applications. The advent of the computer, first thought to be fatal to the subject, turned out to be a stimulus, first of all because it allowed more detailed computations requiring special numerical algorithms, but mainly because it led to automatic summation routines, notably the WZ-method (see Koepf's contribution). So OPSF is a very lively branch of mathematics.

Since more advanced courses on OPSF seldom appear in the curriculum, we felt the need for such courses for young researchers (graduate students and post-docs). A series of European summer schools was started with one in Laredo, Spain (2000) and one in Inzell, Germany (2001). This book contains the notes for the lectures of the summer school in Belgium in 2002, which took place from August 12–16, 2002, at the Katholieke Universiteit Leuven, Belgium. In 2003 a summer school in OPSF will be held in Coimbra, Portugal.

As is clear from the previous paragraphs, there are many different aspects of OPSF and the Leuven summer school focused on computer algebra, representation theory and harmonic analysis, combinatorics, and asymptotics. The

relation between *computer algebra* and special functions was revolutionised by the introduction of very clever algorithms that allow to decide, e.g., for summability of hypergeometric series (Gosper's algorithm, Zeilberger's algorithm, WZ-method, etc.). This makes computer algebra a very important tool in research involving special functions but also a valuable source of research within computer science. The relation between *representation theory* of groups, *harmonic analysis* and special functions is approximately fifty years old, and hence relatively young. The interaction has turned out to be very fruitful on both sides, and it is still developing rapidly, in particular because of its applications in physics, e.g., Racah-Wigner theory of angular momentum, integrable systems (Calogero-Moser-Sutherland), quantum groups and basic hypergeometric series, and Knizhnik-Zamolodchikov equations. One relation between *combinatorics* and special functions is via enumeration, and typical results are the famous Rogers-Ramanujan identities and other identities for partitions of integers. In this field there are many open problems that can be formulated in an elementary way. *Asymptotics*, and related error estimates, are very important in order to describe phenomena for large time or for a large number of degrees of freedom. The classical asymptotic expansions for special functions have recently greatly been improved by allowing exponentially small terms, leading to exponetial asymptotics and hyperasymptotics. Sometimes one obtains asymptotics from an integral representation, or from a differential equation. Another recent development is that boundary value problems can be used, and a Riemann-Hilbert approach combined with a steepest descent method then allows to find uniform asymptotics. There were six series of lectures each of six hours.

Wolfram Koepf discusses the interaction between computer algebra and special functions. The automatic summation algorithms of Gosper, Wilf-Zeilberger and Petkovšek are discussed. Also algorithms for definite and indefinite integration, obtaining generating functions, obtaining hypergeometric expressions, solving recurrence relations and differential, difference and q-difference equations are discussed. This subject is easiest understood during hands-on sessions, as was the case during the Leuven summer school, and to make this possible for the reader of this book the Maple worksheets, including references to other lectures, are available, see Koepf's contribution for the web address. Quite a few of the identities appearing in the other lectures can be obtained using the software described in Koepfs's contribution.

Joris Van der Jeugt discusses the link between Clebsch-Gordan and Racah coefficients for 2 and 3-fold tensor products of simple Lie algebras ($\mathfrak{su}(2)$ and $\mathfrak{su}(1,1)$) and orthogonal polynomials of hypergeometric type, in particular the Hahn and Racah polynomials. He extends this to multivariable orthogonal polynomials by going to n-fold tensor products related to combinatorics on rooted trees. This is closely related to the Racah-Wigner algebra in the theory of angular momentum.

Margit Rösler discusses the Dunkl transform, which is a non-trivial multidimensional generalisation of the one-dimensional Fourier and Hankel trans-

forms. The definition of the Dunkl kernel is in terms of finite reflection groups, and for particular choices of the parameters this transform has a group theoretical interpretation as spherical Fourier transform. Many familiar features of the Fourier transform, such as inversion formula, L^2-theory, generalised Hermite functions as eigenfunctions, positivity of the kernel, asymptotic behaviour of the kernel, have analogues for the Dunkl transform, and are discussed. She describes the analogues of the Laplacian, heat equation, heat semi-group, and the link to Calogero-Moser-Sutherland models (n-particle systems).

Dennis Stanton discusses the interaction between combinatorics, enumeration, additive number theory, and special functions. He uses the q-binomial theorem to derive Ramanujan's congruences for the partition function and other important identities such as the Jacobi triple product identity and the Rogers-Ramanujan identities. Unimodality of the q-binomial coefficient is proved using representation theory as described in the contribution of Joris Van der Jeugt and the Macdonald identity of type B_2 is proved using the Weyl group, which is described in Margit Rösler's contribution. A combinatorial interpretation of the three-term recurrence relation is given for some sets of orthogonal polynomials using Motzkin paths, which allows a combinatorial interpretation of moments, Hankel determinants, and continued fractions. Some open problems, like the Borwein conjecture which is related to the representation theory of the Virasoro algebra, are presented as well.

Arno Kuijlaars discusses asymptotics of orthogonal polynomials using the so-called Riemann-Hilbert method. This method characterises orthogonal polynomials and their Cauchy transforms in terms of matrix valued analytic functions having a jump over a system of contours, typically the real line or an interval. This is a very strong method that has arisen from recent work of Fokas, Its and Kitaev on isomonodromy problems in $2D$ quantum gravity. Deformation of contours combined with a steepest descent method for oscillatory Riemann-Hilbert problems, which was developed by Deift and Zhou for the analysis of the MKdV equation, gives a very powerful tool for obtaining asymptotics for orthogonal polynomials.

Adri Olde Daalhuis discusses asymptotics of functions defined by integrals or as solutions of differential equations. He shows how re-expansions of divergent asymptotic series can be used to obtain exponentially improved asymptotics, both locally and globally. He also discusses the notion of Stokes multipliers and Stokes lines, and he shows how the Stokes lines can be determined from the saddle point method, and how to compute the Stokes multipliers with great accuracy from asymptotic expansions. This method is worked out in detail in several examples.

The lecture notes are aimed at graduate students and post-docs, or anyone who wants to have an introduction to (and learn about) the subjects mentioned. Each of the contributions is self-contained, and contains up to date references to the literature so that anyone who wants to apply the results to his own advantage has a good starting point. The knowledge required for

the lectures is (real and complex) analysis, some basic notions of algebra and discrete mathematics, and some elementary facts of orthogonal polynomials. A computer equipped with Maple software is useful for the lecture related to computer algebra. Exercises are supplied in each of the contributions, and some open problems are discussed in most of them. An extensive index of keywords at the end will be useful for locating the topics of interest. So having mastered the lecture notes gives a good level to read research papers in this field, and to start doing research as well. This has been one of the main scientific goals of the summer school, another main goal being to enhance the interaction between young researchers from various European countries.

The summer school in Orthogonal Polynomials and Special Functions in Leuven has been attended by 60 participants, most of whom are at the beginning of their research. The following institutions have supported the Leuven summer school financially or otherwise: Fonds voor Wetenschappelijk Onderzoek – Vlaanderen (Belgium), the Netherlands Organisation for Scientific Research (NWO), FWO Research Network WO.011.96N "Fundamentele Methoden en Technieken in de Wiskunde" (Belgium), Thomas Stieltjes Institute for Mathematics (the Netherlands), Stichting Computer Algebra Nederland (the Netherlands), Katholieke Universiteit Leuven (Belgium), SIAM Activity Group on Orthogonal Polynomials and Special Functions. The summer school is also part of the Socrates/Erasmus Intensive Programme *Orthogonal Polynomials and Special Functions* of the European Union (29242-IC-1-2001-PT-ERASMUS-IP-13). We thank all these organisations for their support. We thank the lecturers for giving excellent lectures and for preparing the contributions in this volume. We also thank Eric Opdam for his generous support.

Delft and Leuven,
October 2002

Erik Koelink
Walter Van Assche

Contents

Computer Algebra Algorithms for Orthogonal Polynomials and Special Functions
Wolfram Koepf 1
1 The computation of power series and hypergeometric functions 2
2 Summation of hypergeometric series 6
3 Hypergeometric term solutions of recurrence equations 14
4 Integration 15
5 Applications and further algorithms 18
6 Epilogue 22
References 23

$3nj$-Coefficients and Orthogonal Polynomials of Hypergeometric Type
Joris Van der Jeugt 25
1 The Lie algebras $\mathfrak{su}(2)$ and $\mathfrak{su}(1,1)$ and representations 27
2 Hypergeometric series and some transformation formulas 35
3 Tensor product of two $\mathfrak{su}(2)$ representations 40
4 Tensor product of three $\mathfrak{su}(2)$ representations and Racah coefficients 56
5 Tensor product of $n+1$ representations and $3nj$-coefficients 70
References 89

Dunkl Operators: Theory and Applications
Margit Rösler 93
1 Introduction 94
2 Dunkl operators and the Dunkl transform 95
3 CMS models and generalized Hermite polynomials 110
4 Positivity results 117
5 Asymptotic analysis for the Dunkl kernel 127
6 Notation 132
References 132

Enumeration and Special Functions
Dennis Stanton 137
1 Enumeration and q-series 138
2 Orthogonal polynomials 155
3 Symmetric functions 160
References 165

Riemann-Hilbert Analysis for Orthogonal Polynomials
Arno B.J. Kuijlaars 167
1 Introduction 168
2 Boundary values of analytic functions 169
3 Matrix Riemann-Hilbert problems 172
4 Riemann–Hilbert problem for orthogonal polynomials on the real line 177
5 Riemann–Hilbert problem for orthogonal polynomials on $[-1,1]$ . . . 179
6 Basic idea of steepest descent method 181
7 First transformation $Y \mapsto T$ 182
8 Second transformation $T \mapsto S$ 184
9 Special case $\alpha = \beta = -\frac{1}{2}$ 187
10 Model Riemann Hilbert problem 189
11 Third transformation $S \mapsto R$ 191
12 Asymptotics for orthogonal polynomials (case $\alpha = \beta = -\frac{1}{2}$) 192
13 Case of general α and β 196
14 Construction of the local parametrix 198
15 Asymptotics for orthogonal polynomials (general case) 204
References 208

Exponential Asymptotics
Adri B. Olde Daalhuis 211
1 Introduction 212
2 Local hyperasymptotics for integrals 217
3 Global hyperasymptotics for integrals 221
4 Global hyperasymptotics for linear ODEs 226
5 Exercises 241
References 243

Index 245

Computer Algebra Algorithms for Orthogonal Polynomials and Special Functions

Wolfram Koepf

Department of Mathematics and Computer Science, University of Kassel, Heinrich-Plett-Str. 40, D–34132 Kassel, Germany,
`koepf@mathematik.uni-kassel.de`

Summary. In this minicourse I would like to present computer algebra algorithms for the work with orthogonal polynomials and special functions. This includes

- the computation of power series representations of hypergeometric type functions, given by "expressions", like $\arcsin(x)/x$,
- the computation of holonomic differential equations for functions, given by expressions,
- the computation of holonomic recurrence equations for sequences, given by expressions, like $\binom{n}{k}\frac{x^k}{k!}$,
- the identification of hypergeometric functions,
- the computation of antidifferences of hypergeometric terms (Gosper's algorithm),
- the computation of holonomic differential and recurrence equations for hypergeometric series, given the series summand, like
$$P_n(x) = \sum_{k=0}^{n} \binom{n}{k} \binom{-n-1}{k} \left(\frac{1-x}{2}\right)^k$$
(Zeilberger's algorithm),
- the computation of hypergeometric term representations of series (Zeilberger's and Petkovšek's algorithm),
- the verification of identities for (holonomic) special functions,
- the detection of identities for orthogonal polynomials and special functions,
- the computation with Rodrigues formulas,
- the computation with generating functions,
- corresponding algorithms for q-hypergeometric (basic hypergeometric) functions,
- the identification of classical orthogonal polynomials, given by recurrence equations.

All topics are properly introduced, the algorithms are discussed in some detail and many examples are demonstrated by Maple implementations. In the lecture, the participants are invited to submit and compute their own examples.

Let us remark that as a general reference we use the book [11], the computer algebra system Maple [16], [4] and the Maple packages **FPS** [9], [7], **gfun** [19], **hsum** [11], **infhsum** [22], **hsols** [21], **qsum** [2] and **retode** [13].

E. Koelink and W. Van Assche (Eds.): LNM 1817, pp. 1–24, 2003.

1 The computation of power series and hypergeometric functions 2
1.1 Hypergeometric series 3
1.2 Holonomic differential equations 4
1.3 Algebra of holonomic functions 4
1.4 Hypergeometric power series 5
1.5 Identification of hypergeometric functions 5
2 Summation of hypergeometric series 6
2.1 Fasenmyer's method 6
2.2 Indefinite summation and Gosper's algorithm 7
2.3 Zeilberger's algorithm 8
2.4 A generating function problem 11
2.5 Automatic computation of infinite sums 11
2.6 The WZ method 12
2.7 Differential equations 13
3 Hypergeometric term solutions of recurrence equations . . . 14
3.1 Petkovšek's algorithm 14
3.2 Combining Zeilberger's and Petkovšek's algorithms 14
4 Integration 15
4.1 Indefinite integration 15
4.2 Definite integration 15
4.3 Rodrigues formulas 17
4.4 Generating functions 17
5 Applications and further algorithms 18
5.1 Parameter derivatives 18
5.2 Basic hypergeometric summation 19
5.3 Orthogonal polynomial solutions of recurrence equations 20
6 Epilogue 22
References 23

1 The computation of power series and hypergeometric functions

Given an expression $f(x)$ in the variable x, one would like to find the Taylor series

$$f(x) = \sum_{k=0}^{\infty} a_k\, x^k \ ,$$

i.e., a formula for the coefficient a_k. For example, if $f(x) = \exp(x)$, then

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k ,$$

hence $a_k = \frac{1}{k!}$. If the result is simple enough, the `FPS` (formal power series) procedure of the Maple package `FPS.mpl` ([9], [7]) computes this series, even if it is a Laurent series (including negative powers) or Puiseux series (including rational powers).

The main idea behind this procedure is

1. to compute a differential equation for $f(x)$,
2. to convert the differential equation to a recurrence equation for a_k,
3. and to solve the recurrence equation for a_k.

1.1 Hypergeometric series

The above procedure is successful at least if $f(x)$ is hypergeometric. A series

$$\sum_{k=0}^{\infty} a_k$$

is called hypergeometric, if the series coefficient a_k has rational term ratio

$$\frac{a_{k+1}}{a_k} \in \mathbb{Q}(k) .$$

The function

$${}_pF_q\left(\left.\begin{matrix} a_1, a_2, \ldots, a_p \\ b_1, b_2, \ldots, b_q \end{matrix}\right| x\right) := \sum_{k=0}^{\infty} A_k\, x^k = \sum_{k=0}^{\infty} \frac{(a_1)_k (a_2)_k \cdots (a_p)_k}{(b_1)_k (b_2)_k \cdots (b_q)_k} \frac{x^k}{k!} \tag{1.1}$$

is called the *generalized hypergeometric series*, since its term ratio

$$\frac{A_{k+1}\, x^{k+1}}{A_k\, x^k} = \frac{(k+a_1)\cdots(k+a_p)}{(k+b_1)\cdots(k+b_q)} \frac{x}{(k+1)} \tag{1.2}$$

is a general rational function, in factorized form. Here $(a)_k = a(a+1)\cdots(a+k-1)$ denotes the *Pochhammer symbol* or shifted factorial. The summand a_k of the generalized hypergeometric series is called a *hypergeometric term*.

The Maple commands `factorial` (short form `!`), `pochhammer`, `binomial`, and `GAMMA` can be used to enter the corresponding functions, `hypergeom` denotes the hypergeometric series, and the `hyperterm` command of the **sumtools** and **hsum** packages denotes a hypergeometric term.[1]

[1] The package `sumtools` is part of Maple [4]. Note that Maple 8 contains a second package `SumTools` ([15], [4]) which also contains summation algorithms.

1.2 Holonomic differential equations

A homogeneous linear differential equation with polynomial coefficients is called *holonomic*. If $f(x)$ satisfies a holonomic differential equation, then its Taylor series coefficients a_k satisfy a holonomic recurrence equation, and vice versa.

To find a holonomic differential equation for an expression $f(x)$, one differentiates $f(x)$, and writes the sum

$$\sum_{j=0}^{J} c_j f^{(j)}(x)$$

as a sum of (over $\mathbb{Q}(x)$) linearly independent summands, whose coefficients should be zero. This gives a system of linear equations for $c_j \in \mathbb{Q}(x)$ $(j = 0, \ldots, J)$. If it has a solution, we have found a differential equation with rational function coefficients, and multiplying by their common denominator yields the equation sought for.

Iterating this procedure for $J = 1, 2, \ldots$ yields the holonomic differential equation of lowest order valid for $f(x)$.

The command `HolonomicDE`[2] of the **FPS** package is an implementation of this algorithm.

Exercise 1. Find a holonomic differential equation for $f(x) = \sin(x)\exp(x)$. Use the algorithm described. Don't use the **FPS** package.
Using the **FPS** package `FPS.mpl`, find a holonomic differential equation for $f(x)$ and for $g(x) = \arcsin(x)^3$.

1.3 Algebra of holonomic functions

A function that satisfies a holonomic differential equation is called a holonomic function. Sum and product of holonomic functions turn out to be holonomic, and their representing differential equations can be computed from the differential equations of the summands and factors, respectively, by linear algebra.

We call a sequence that satisfies a holonomic recurrence equation a holonomic sequence. Sum and product of holonomic sequences are holonomic, and similar algorithms exist. As already mentioned, a function is holonomic if and only if it is the generating function of a holonomic sequence.

The **gfun** package by Salvy and Zimmermann [19] contains—besides others—implementations of the above algorithms.

[2] In earlier versions of the **FPS** package the command name was `SimpleDE`.

Exercise 2. Use the **gfun** package to generate differential equations for $f(x) = \sin(x)\exp(x)$, and $g(x) = \sin(x) + \exp(x)$ by utilizing the (known) ODEs for the summands and factors, respectively.
Use the **gfun** package to generate recurrence equations for

$$a_k = k\binom{n}{k}^2 \quad \text{and} \quad b_k = k + \binom{n}{k}^2 .$$

1.4 Hypergeometric power series

Having found a holonomic differential equation for $f(x)$, by substituting

$$f(x) = \sum_{k=0}^{\infty} a_k\, x^k \ ,$$

and equating coefficients, it is easy to deduce a holonomic recurrence equation for a_k.

If we are lucky, the recurrence is of first order, hence the function is a hypergeometric series, and the coefficients can be computed by (1.1)–(1.2).

The command `SimpleRE` of the **FPS** package combines the above steps and computes a recurrence equation for the series coefficients of an expression.

1.5 Identification of hypergeometric functions

Assume, we have

$$F = \sum_{k=0}^{\infty} a_k \ .$$

How do we find out which ${}_pF_q(x)$ this is?

The simple idea is to write the ratio $\frac{a_{k+1}}{a_k}$ as factorized rational function, and to read off the upper and lower parameters according to (1.2).

The command `Sumtohyper` of the **sumtools** and **hsum** packages are implementations of this algorithm.

Exercise 3. Write $\cos(x)$ in hypergeometric notation by hand computation. Use the **sumtools** package to do the same. Restart your session and use the **hsum** package `hsum6.mpl` instead.
Get the hypergeometric representations for $\sin(x)$, $\sin(x)^2$, $\arcsin(x)$, $\arcsin(x)^2$, and $\arctan(x)$, combining **FPS** and **hsum**.

Exercise 4. Write the following representations of the *Legendre polynomials* in hypergeometric notation:

$$P_n(x) = \sum_{k=0}^{n} \binom{n}{k} \binom{-n-1}{k} \left(\frac{1-x}{2}\right)^k \tag{1.3}$$

$$= \frac{1}{2^n} \sum_{k=0}^{n} \binom{n}{k}^2 (x-1)^{n-k} (x+1)^k \tag{1.4}$$

$$= \frac{1}{2^n} \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \binom{n}{k} \binom{2n-2k}{n} x^{n-2k} \,. \tag{1.5}$$

In the hypergeometric representations, where from can you read off the upper bound of the sum?

2 Summation of hypergeometric series

In this section, we try to simplify both definite and indefinite hypergeometric series.

2.1 Fasenmyer's method

Given a sequence s_n, as hypergeometric sum

$$s_n = \sum_{k=-\infty}^{\infty} F(n,k) \,,$$

how do we find a recurrence equation for s_n? Celine Fasenmyer proposed the following algorithm (see e.g., [11], Chapter 4):

1. Compute ansatz $:= \sum\limits_{\substack{i=0,\ldots,I \\ j=0,\ldots,J}} \dfrac{F(n+j,k+i)}{F(n,k)} \in \mathbb{Q}(n,k)$.
2. Bring this into rational form and set the numerator coefficient list w.r.t. k zero. If the corresponding linear system has a solution, this leads to a k-free recurrence equation for the summand $F(n,k)$.
3. Summing this recurrence equation for $k=-\infty,\ldots,\infty$ gives the desired recurrence for s_n.

If successful, this results in a holonomic recurrence equation for s_n. If we are lucky, and the recurrence is of first order, then the sum can be written as a hypergeometric term by formula (1.1)–(1.2). This algorithm can be accessed by the commands `kfreerec` and `fasenmyer` of the **hsum** package.

As an example, to compute

$$s_n = \sum_{k=0}^{n} F(n,k) = \sum_{k=0}^{n} \binom{n}{k} ,$$

in the first step one gets the well-known binomial coefficient recurrence

$$F(n+1,k) = F(n,k) + F(n,k-1)$$

or in the usual notation

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1} ,$$

from which it follows by summation for $k = -\infty, \ldots, \infty$

$$s_{n+1} = s_n + s_n = 2s_n ,$$

since

$$\sum_{k=-\infty}^{\infty} F(n,k) = \sum_{k=-\infty}^{\infty} F(n,k-1) .$$

With $s_0 = 1$ one finally gets $s_n = 2^n$.

In practice, however, Fasenmyer's algorithm is rather slow and inefficient.

Exercise 5. Using Fasenmyer's method, compute a three-term recurrence equation for the *Laguerre polynomials*

$$L_n(x) = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \binom{n}{k} x^k = {}_1F_1\left(\left.\begin{array}{c} -n \\ 1 \end{array}\right| x\right)$$

and for the *generalized Laguerre polynomials*

$$L_n^{(\alpha)}(x) = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \binom{n+\alpha}{n-k} x^k. \tag{2.1}$$

2.2 Indefinite summation and Gosper's algorithm

Given a sequence a_k, one would like to find a sequence s_k which satisfies

$$a_k = s_{k+1} - s_k = \Delta s_k . \tag{2.2}$$

Having found s_k makes definite summation easy since by telescoping it follows from (2.2) for arbitrary $M, N \in \mathbb{Z}$

$$\sum_{k=M}^{N} a_k = s_{N+1} - s_M .$$

We call $s_k = \sum a_k$ an indefinite sum (or an antidifference) of a_k. Hence indefinite summation is the inverse of the forward difference operator Δ.

Gosper's algorithm ([6], see e.g., [11], Chapter 5) takes a hypergeometric term a_k and decides whether or not a_k has a hypergeometric term antidifference, and computes it in the affirmative case. In the latter case s_k is a rational multiple of a_k, $s_k = R_k\, a_k$ with $R_k \in \mathbb{Q}(k)$.

Note that whenever Gosper's algorithm does not find a hypergeometric term antidifference, it has therefore *proved* that no such antidifference exists. In particular, using this approach, it is easily shown that the *harmonic numbers* $H_n = \sum\limits_{k=1}^{n} \frac{1}{k}$ cannot be written as a hypergeometric term. On the other hand, one gets (checking the result applying Δ is easy!)

$$\sum a_k = \sum (-1)^k \binom{n}{k} = -\frac{k}{n}\, a_k \ .$$

Both Maple's **sumtools** package and `hsum6.mpl` contain an implementation of Gosper's algorithm by the author. The `gosper` command of the **hsum** package will give error messages that let the user know whether the input is not a hypergeometric term (and hence the algorithm is not applicable) or whether the algorithm has deduced that no hypergeometric term antidifference exists. Since this is (unfortunately) against Maple's general policy, `sumtools[gosper]` does not do so, and gives `FAIL` in these cases.

Exercise 6. Use Gosper's algorithm to compute

$$s(m,n) = \sum_{k=0}^{m} (-1)^k \binom{n}{k} \ ,$$

$$t_n = \sum_{k=1}^{n} k^3 \ ,$$

and

$$u_n = \sum_{k=1}^{n} \frac{1}{k(k+5)} \ .$$

2.3 Zeilberger's algorithm

Zeilberger [24] had the brilliant idea to use a modified version of Gosper's algorithm to compute definite hypergeometric sums

$$s_n = \sum_{k=-\infty}^{\infty} F(n,k) \ ,$$

like Fasenmyer's algorithm does (see e.g., [11], Chapter 7). However, Zeilberger's algorithm is much more efficient than Fasenmyer's. Note that, whenever s_n is itself a hypergeometric term, then Gosper's algorithm, applied to $F(n,k)$, fails! Thus a direct application of Gosper's algorithm to the summand is not possible.

Zeilberger's algorithm works as follows:

1. For suitable $J \in \mathbb{N}$ set
$$a_k = F(n,k) + \sigma_1 F(n+1,k) + \cdots + \sigma_J F(n+J,k) \ .$$
2. Apply Gosper's algorithm to determine a hypergeometric term s_k, and at the same time rational functions $\sigma_j \in \mathbb{Q}(n)$ such that $a_k = s_{k+1} - s_k$.
3. Summing for $k = -\infty, \ldots, \infty$ yields the desired holonomic recurrence equation for s_n by telescoping.

One can prove that Zeilberger's algorithm terminates for suitable input.

The command `sumtools[sumrecursion]` as well as `sumrecursion` and `closedform` of the **hsum** package are implementations of Zeilberger's algorithm.

Exercise 7. Compute recurrence equations for the binomial power sums
$$\sum_{k=0}^{n} \binom{n}{k}^m$$
for $m = 2, \ldots, 7$.
In the 1980s these results were worth a paper in a mathematical journal!

If the resulting recurrence equation is of first order, then in combination with formula (1.1)–(1.2) and the value s_0 one gets a hypergeometric term representation of the sum s_n. This is the strategy of the `closedform` command.

As an example, using this approach, it is easy to deduce *Dougall's identity*
$$
{}_7F_6\left(\left.\begin{array}{c} a, 1+\frac{a}{2}, b, c, d, 1+2a-b-c-d+n, -n \\ \frac{a}{2}, 1+a-b, 1+a-c, 1+a-d, b+c+d-a-n, 1+a+n \end{array}\right| 1\right)
$$
$$
= \frac{(1+a)_n\,(a+1-b-c)_n\,(a+1-b-d)_n\,(a+1-c-d)_n}{(1+a-b)_n\,(1+a-c)_n\,(1+a-d)_n\,(1+a-b-c-d)_n}
$$
from its left hand side.

Exercise 8. Find hypergeometric term representations for the sums

$$s_n = \sum_{k=0}^{n} k \binom{n}{k} ,$$

$$t_n = \sum_{k=0}^{n} \binom{n}{k}^2 ,$$

$$u_n = \sum_{k=0}^{n} (-1)^k \binom{n}{k}^2 ,$$

and

$$v_n = \sum_{k=0}^{n} \binom{a}{k} \binom{b}{n-k} .$$

Assume you have two hypergeometric sums, how can you check whether they are different disguised versions of the same special function? Zeilberger's paradigm is to compute their recurrence equations, and, if these agree, then it remains to check enough initial values. Using this approach, it is easy to check that the three representations of the Legendre polynomials, given in (1.3)–(1.5), all agree with the fourth representation

$$P_n(x) = x^n \, {}_2F_1\left(\left. \begin{matrix} \frac{n}{2}, \frac{1-n}{2} \\ 1 \end{matrix} \right| 1 - \frac{1}{x^2} \right) \tag{2.3}$$

Exercise 9. Prove that the representations (1.3)–(1.5) and (2.3) all constitute the same functions, the Legendre polynomials.

With this method, one can prove many of the hypergeometric identities that appear in Joris van der Jeugt's contribution in these lecture notes, for example Whipple's transformation (2.10) as well as the different representations of the Clebsch-Gordan coefficients `Clebsch-Gordan coefficients` and Racah polynomials, and many of the corresponding exercises can be solved automatically.

Even more advanced questions that involve double sums can be solved: to prove *Clausen's formula*

$${}_2F_1\left(\left. \begin{matrix} a, b \\ a + b + \frac{1}{2} \end{matrix} \right| x \right)^2 = {}_3F_2\left(\left. \begin{matrix} 2a, 2b, a + b \\ 2a + 2b, a + b + \frac{1}{2} \end{matrix} \right| x \right) ,$$

which gives the cases when a ${}_3F_2$ is the square of a ${}_2F_1$, one can write the left hand side as a Cauchy product. This gives a double sum. It turns out that the inner sum can be simplified by Zeilberger's algorithm, and the remaining sum is exactly the right hand side.

2.4 A generating function problem

Recently, Folkmar Bornemann showed me a generating function of the Legendre polynomials and asked me to generate it automatically [3]. Here is the problem: write

$$G(x,z,\alpha) := \sum_{n=0}^{\infty} \binom{\alpha+n-1}{n} P_n(x)\, z^n$$

as a hypergeometric function. We can take any of the four given hypergeometric representations of the Legendre polynomials to write $G(x,z,\alpha)$ as double sum. Then the trick is to change the order of summation. If we are lucky, then the inner sum is Zeilberger-summable, hence a single hypergeometric sum remains which gives the desired result.

It turns out that (only) the fourth representation (2.3) leads to such a result, namely

$$\sum_{n=0}^{\infty} \binom{\alpha+n-1}{n} P_n(x)\, z^n = \frac{1}{(1-xz)^\alpha}\, {}_2F_1\left(\left.\begin{matrix} \frac{\alpha}{2}, \frac{\alpha+1}{2} \\ 1 \end{matrix}\right| \frac{(x^2-1)\, z^2}{(xz-1)^2}\right) .$$

Note that a variant of this identity can be found as number 05.03.23.0006.01 in the extensive online handbook [23]. It occurs there without citation, though, hence without proof.

Advanced Exercise 10. Derive the *Askey-Gasper identity* which was the essential ingredient in de Branges' proof of the Bieberbach conjecture (see, e.g., [11], Example 7.4). Hence write as a single hypergeometric series

$$\sum_{j=0}^{\lfloor \frac{n-k}{2} \rfloor} \frac{\left(\frac{1}{2}\right)_j \left(\frac{\alpha}{2}+1\right)_{n-j} \left(\frac{\alpha+3}{2}\right)_{n-2j} (\alpha+1)_{n-2j}}{j!\, \left(\frac{\alpha+3}{2}\right)_{n-j} \left(\frac{\alpha+1}{2}\right)_{n-2j} (n-2j)!}$$
$$\qquad {}_3F_2\left(\left.\begin{matrix} 2j-n, n-2j+\alpha+1, \frac{\alpha+1}{2} \\ \alpha+1, \frac{\alpha+2}{2} \end{matrix}\right| x\right).$$

2.5 Automatic computation of infinite sums

Whereas Zeilberger's algorithm finds *Chu-Vandermonde's formula* for $n \in \mathbb{N}_{\geq 0}$

$${}_2F_1\left(\left.\begin{matrix} -n, b \\ c \end{matrix}\right| 1\right) = \frac{(c-b)_n}{(c)_n} , \tag{2.4}$$

the question arises how to detect *Gauss' identity*

$$ {}_2F_1\left(\left.\begin{array}{c} a, b \\ c \end{array}\right| 1\right) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} $$

for $a, b, c \in \mathbb{C}$ in case of convergence, i.e., if Re $(c-a-b) > 0$, extending (2.4).

The idea is to detect by Zeilberger's algorithm

$$ {}_2F_1\left(\left.\begin{array}{c} a, b \\ c+m \end{array}\right| 1\right) = \frac{(c-a)_m\,(c-b)_m}{(c)_m\,(c-a-b)_m}\, {}_2F_1\left(\left.\begin{array}{c} a, b \\ c \end{array}\right| 1\right) $$

and then consider the limit as $m \to \infty$.

Using appropriate limits for the Gamma function, this and similar questions can be handled automatically by the `infclosedform` prodecure of Vidunas' and Koornwinder's Maple package **infhsum** [22].

Exercise 11. Derive *Kummer's theorem*, i.e. find a closed form for

$$ {}_2F_1\left(\left.\begin{array}{c} a, b \\ 1+a-b \end{array}\right| -1\right) . $$

Find an extension of the *Pfaff-Saalschütz formula*

$$ {}_3F_2\left(\left.\begin{array}{c} a, b, -n \\ c, 1+a+b-c-n \end{array}\right| 1\right) = \frac{(c-a)_n\,(c-b)_n}{(c)_n\,(c-a-b)_n} . $$

2.6 The WZ method

Assume we want to prove an identity

$$ \sum_{k=-\infty}^{\infty} f(n,k) = \tilde{s}_n $$

with hypergeometric terms $f(n,k)$ and $\tilde{s}_n$, see e.g. [11], Chapter 6. Then Wilf's idea is to divide by $\tilde{s}_n$, and therefore to put the identity into the form

$$ s_n := \sum_{k=-\infty}^{\infty} F(n,k) = 1 . \tag{2.5} $$

Now we can apply Gosper's algorithm to $F(n+1,k) - F(n,k)$ as a function of k. If this is successful, then it generates a rational multiple $G(n,k)$ of $F(n,k)$, i.e., $G(n,k) = R(n,k)\,F(n,k)$ with $R(n,k) \in \mathbb{Q}(n,k)$, such that

$$ F(n+1,k) - F(n,k) = G(n,k+1) - G(n,k) , \tag{2.6} $$

and telescoping yields $s_{n+1} - s_n = 0$, and therefore (2.5). Hence the sheer success of Gosper's algorithm gives a proof of (2.5). This is called the WZ method.

Moreover, if the WZ method was successful, it has computed the *rational certificate* $R(n,k) \in \mathbb{Q}(n,k)$ which enables a completely independent proof of (2.5) that can be carried out by hand computations: Dividing (2.6) by $F(n,k)$, we have only to prove

$$\frac{F(n+1,k)}{F(n,k)} - 1 = R(n,k+1)\frac{F(n,k+1)}{F(n,k)} - R(n,k) \ ,$$

a purely rational identity.

The function `WZcertificate` of the **hsum** package computes the rational certificate if applicable.

Exercise 12. Prove by the WZ method:

$$\sum_{k=1}^{n} k\binom{n}{k} = n2^{n-1}$$

and (this is again the disguised Chu-Vandermonde formula)

$$\sum_{k=0}^{n} \binom{a}{k}\binom{b}{n-k} = \binom{a+b}{n} \ .$$

2.7 Differential equations

Zeilberger's algorithm can easily be adapted to generate holonomic differential equations for *hyperexponential sums* (see, e.g., [11], Chapter 10)

$$s(x) = \sum_{k=-\infty}^{\infty} F(x,k) \ .$$

For this purpose, the summand $F(x,k)$ must be a hyperexponential term w.r.t. x:

$$\frac{\frac{\partial}{\partial x}F(x,k)}{F(x,k)} \in \mathbb{Q}(x,k) \ .$$

With this algorithm which is implemented as `sumdiffeq` in the **hsum** package it is easy to check that all representations (1.3)–(1.5) and (2.3) of the Legendre polynomials satisfy the same differential equation

$$(1-x^2)\,P_n''(x) - 2x\,P_n'(x) + n(n+1)\,P_n(x) = 0 \ .$$

In CAOP [20], an online version of the *Askey-Wilson scheme* of orthogonal polynomials [8] developed by René Swarttouw, the **hsum** package is used

to interactively compute recurrence and differential equations for personally standardized orthogonal polynomial families of the Askey-Wilson scheme.

Exercise 13. Find holonomic differential equations for the *Jacobi polynomials*

$$P_n^{(\alpha,\beta)}(x) = \binom{n+\alpha}{n} {}_2F_1\left(\left.\begin{matrix} -n, n+\alpha \\ \alpha+\beta+1 \end{matrix}\right| \frac{1+x}{2}\right), \tag{2.7}$$

and the generalized Laguerre polynomials (2.1).

3 Hypergeometric term solutions of recurrence equations

3.1 Petkovšek's algorithm

Petkovšek's algorithm is an adaption of Gosper's (see, e.g., [11], Chapter 9). Given a holonomic recurrence equation, it determines all hypergeometric term solutions. The command `rechyper` of the **hsum** package is an implementation of Petkovšek's algorithm.

Petkovšek's algorithm is slow, especially if the leading and trailing coefficients of the recurrence equation have many factors. Maple 9 will contain a much more efficient algorithm **hsols** due to Mark van Hoeij [21] for the same purpose.

As an example, the recurrence equation

$$3(3n+4)(3n+7)(3n+8)s_{n+3} + 4(3n+4)(37n^2+180n+218)s_{n+2}$$

$$+16(n+2)(33n^2+125n+107)s_{n+1} + 64(n+1)(n+2)(3n+7)s_n = 0$$

which is the output of Zeilberger's algorithm applied to the sum

$$s_n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \binom{4k}{n},$$

has the hypergeometric term solution $(-4)^n$ which finally yields $s_n = (-4)^n$.

3.2 Combining Zeilberger's and Petkovšek's algorithms

As seen, Zeilberger's algorithm may not give a recurrence equation of first order, even if the sum is a hypergeometric term. This rarely happens, though. In such a case, the combination of Zeilberger's with Petkovšek's algorithm guarantees to find out whether a given sum can be written as a hypergeometric term.

Exercise 9.3 of my book [11] gives 9 examples for this situation, all from p. 556 of [18].

Advanced Exercise 14. Use a combination of Zeilberger's algorithm and Petkovšek's algorithm to find a simple representation (as linear combination of two hypergeometric terms) of the sum

$$s_n = \sum_{k=0}^{\lfloor n/3 \rfloor} \binom{n-2k}{k} \left(-\frac{4}{27}\right)^k .$$

4 Integration

4.1 Indefinite integration

To find holonomic recurrence and differential equations for hypergeometric and hyperexponential integrals, one needs a continuous version of Gosper's algorithm. Almkvist and Zeilberger gave such an algorithm ([1], see, e.g., [11], Chapter 11). It finds hyperexponential antiderivatives if those exist. This algorithm is accessible as procedure `contgosper` of the **hsum** package.

For example, this algorithm proves that the function e^{x^2} does not have a hyperexponential antiderivative. In fact, this function does not even have an elementary antiderivative; but this cannot be detected by the given algorithm. On the other hand, the algorithm computes, e.g., the integral

$$\int \left(\frac{2x}{1-x^{10}} + \frac{10(1+x^2)x^9}{(1-x^{10})^2} \right) dx = \frac{1+x^2}{1-x^{10}} .$$

4.2 Definite integration

Applying the continuous Gosper algorithm, one can easily adapt the discrete versions of Zeilberger's algorithm to the continuous case. The resulting algorithms find holonomic recurrence and differential equations for hypergeometric and hyperexponential integrals.

The procedures `intrecursion` and `intdiffeq` of the **hsum** package are implementations of these algorithms, see [11], Chapter 12.

As an example, we would like to find

$$S(x,y) = \int\limits_0^1 t^x \, (1-t)^y \, dt .$$

Applying the continuous Zeilberger algorithm w.r.t. x and y, respectively, results in the two recurrence equations

$$-(x+y+2)\, S(x+1,y) + (x+1)\, S(x) = 0$$

and

$$-(x+y+2)\,S(x,y+1)+(x+1)\,S(x)=0\;.$$

Solving both recurrence equations (e.g., with Maple's `rsolve` command) shows that $S(x,y)$ must be a multiple of

$$\frac{\Gamma(x+1)\,\Gamma(y+1)}{\Gamma(x+y+2)}\;.$$

Computing the initial value

$$S(0,0)=\int_0^1 dt=1$$

proves the identity

$$S(x,y)=\frac{\Gamma(x+1)\,\Gamma(y+1)}{\Gamma(x+y+2)}$$

for $x,y\in\mathbb{Z}$. Since we work with recurrence equations, this method cannot find the result for other complex values x,y.

Another example is given by the integral

$$I(x)=\int_0^\infty \frac{x^2}{(x^4+t^2)(1+t^2)}dt$$

for which the algorithm detects the holonomic differential equation

$$x(x^4-1)\,I''(x)+(1+7x^4)\,I'(x)+8x^3\,I(x)=0\;.$$

Maple's `dsolve` command finds the solution

$$I(x)=\frac{\pi}{2(x^2+1)}\;.$$

Advanced Exercise 15. Write the integral

$$\int_0^1 t^{c-1}\,(1-t)^{d-1}\;{}_2F_1\left(\left.\begin{array}{c}a,b\\c\end{array}\right|tx\right)$$

as a hypergeometric series. This generates the so-called *Bateman integral representation*. For which $c,d\in\mathbb{C}$ is the result valid?

Advanced Exercise 16. Find a similar representation for the integral

$$\int_0^1 t^{c-1}\,(1-t)^{d-1}\;{}_2F_1\left(\left.\begin{array}{c}a,b\\d\end{array}\right|tx\right)\;.$$

4.3 Rodrigues formulas

Using Cauchy's integral formula

$$h^{(n)}(z) = \frac{n!}{2\pi i} \int\limits_{\gamma} \frac{h(t)}{(t-x)^{n+1}} dt$$

for the nth derivative makes the integration algorithms accessible for Rodrigues type expressions

$$f_n(x) = g_n(x) \frac{d^n}{dx^n} h_n(x) \ .$$

This is implemented in`rodriguesrecursion` and `rodriguesdiffeq` of the **hsum** package, see [11], Chapter 13.

Using these algorithms, one can easily show that the functions

$$P_n(x) = \frac{(-1)^n}{2^n \, n!} \frac{d^n}{dx^n} (1-x^2)^n$$

are the Legendre polynomials, and that

$$L_n^{(\alpha)}(x) = \frac{e^x}{n! \, x^\alpha} \frac{d^n}{dx^n} \left(e^{-x} \, x^{\alpha+n} \right)$$

are the generalized Laguerre polynomials.

Exercise 17. Prove the Rodrigues formula

$$P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{2^n \, n!} (1-x)^{-\alpha} \, (1+x)^{-\beta} \frac{d^n}{dx^n} \left((1-x)^\alpha (1+x)^\beta (1-x^2)^n \right)$$

for the Jacobi polynomials (2.7).

4.4 Generating functions

If $F(z)$ is the generating function of the sequence $a_n f_n(x)$,

$$F(z) = \sum_{n=0}^{\infty} a_n \, f_n(x) \, z^n \ ,$$

then by Cauchy's formula and Taylor's theorem

$$f_n(x) = \frac{1}{a_n} \frac{F^{(n)}(0)}{n!} = \frac{1}{a_n} \frac{1}{2\pi i} \int\limits_{\gamma} \frac{F(t)}{t^{n+1}} dt \ .$$

Hence, again, we can apply the integration algorithms. This is implemented in the functions `GFrecursion` and `GFdiffeq` of the **hsum** package, see [11], Chapter 13.

Using these algorithms, we can easily prove the generating function identity

$$(1-z)^{-\alpha-1}\, e^{\frac{xz}{z-1}} = \sum_{n=0}^{\infty} L_n^{(\alpha)}(x)\, z^n$$

for the generalized Laguerre polynomials.

Exercise 18. Prove that

$$\frac{1}{\sqrt{1-2xz+z^2}} = \sum_{n=0}^{\infty} P_n(x)\, z^n$$

is the generating function of the Legendre polynomials.

Advanced Exercise 19. Write the exponential generating function

$$F(x) = \sum_{n=0}^{\infty} \frac{1}{n!}\, P_n(x)\, z^n$$

of the Legendre polynomials in terms of Bessel functions.
Hint: Use one of the hypergeometric representations of the Legendre polynomials and change the order of summation.

5 Applications and further algorithms

5.1 Parameter derivatives

For some applications, one uses parametrized families of orthogonal polynomials like the generalized Laguerre polynomials that are parametrized by the parameter α. It might be necessary to know the rate of change of the family in the direction of the parameter α (see [10]).

Using Zeilberger's algorithm and limit computations (with Maple's `limit`) one can compute such parameter derivatives in this and in similiar occasions.

Advanced Exercise 20. Prove the following representation for the parameter derivative of the generalized Laguerre polynomials

$$\frac{\partial}{\partial\alpha} L_n^{(\alpha)}(x) = \sum_{k=0}^{n-1} \frac{1}{n-k}\, L_k^{(\alpha)}(x)$$

by proving first that

$$L_n^{(\alpha+\mu)}(x) = \sum_{k=0}^{n} \frac{(\mu)_{n-k}}{(n-k)!}\, L_k^{(\alpha)}(x)$$

and taking limit $\mu \to 0$.

5.2 Basic hypergeometric summation

Instead of considering series whose coefficients A_k have rational term ratio $A_{k+1}/A_k \in \mathbb{Q}(k)$, we can also consider such series whose coefficients A_k have term ratio $A_{k+1}/A_k \in \mathbb{Q}(q^k)$ for some $q \in \mathbb{C}$. This leads to the *q-hypergeometric series*—also called basic hypergeometric series—(see, e.g., [5])

$$ {}_r\varphi_s\left(\left.\begin{matrix} a_1, \ldots, a_r \\ b_1, \ldots, b_s \end{matrix}\right| q; x\right) = \sum_{k=0}^{\infty} A_k\, x^k \; . $$

Here the coefficients are given by

$$ A_k = \frac{(a_1;q)_k \cdots (a_r;q)_k}{(b_1;q)_k \cdots (b_s;q)_k} \frac{x^k}{(q;q)_k} \left((-1)^k\, q^{\binom{k}{2}}\right)^{1+s-r} \quad , $$

where

$$ (a;q)_k = \prod_{j=0}^{k-1} (1 - a q^j) $$

denotes the *q-Pochhammer symbol.*

Further q-expressions are given by

1. the infinite q-Pochhammer symbol: $(a;q)_\infty = \lim_{n\to\infty} (a;q)_n$;
2. the q-factorial: $[k]_q! = \dfrac{(q;q)_k}{(1-q)^k}$;
3. the q-Gamma function: $\Gamma_q(z) = \dfrac{(q;q)_\infty}{(q^z;q)_\infty}(1-q)^{1-z}$;
4. the q-binomial coefficient: $\begin{bmatrix} n \\ k \end{bmatrix}_q = \dfrac{(q;q)_n}{(q;q)_k\,(q;q)_{n-k}}$;
5. the q-brackets: $[k]_q = \dfrac{1-q^k}{1-q} = 1 + q + \cdots + q^{k-1}$.

For many of the algorithms mentioned in this minicourse corresponding q-versions exist, see [11]. These are implemented in the **qsum** package, see [2], and the above q-expressions are accessible, see [11], Chapter 3.

For all classical hypergeometric theorems corresponding q-versions exist. These can be proved by a q-version of Zeilberger's algorithm (`qsumrecursion`) via the **qsum** package. For example, the *q-Chu-Vandermonde theorem* states that

$$ {}_2\varphi_1\left(\left.\begin{matrix} q^{-n}, b \\ c \end{matrix}\right| q; \frac{c\,q^n}{b}\right) = \frac{(c/b;q)_n}{(c;q)_n} \; . $$

As usual, the right hand side can be computed from the left hand side.

All classical orthogonal families have q-hypergeometric equivalents. For example, the *little* and the *big q-Legendre polynomials*, respectively, are given by

$$p_n(x|q) = {}_2\varphi_1\left(\begin{matrix} q^{-n}, q^{n+1} \\ q \end{matrix}\,\middle|\, q; qx\right)$$

and

$$P_n(x;c;q) = {}_3\varphi_2\left(\begin{matrix} q^{-n}, q^{n+1}, x \\ q, cq \end{matrix}\,\middle|\, q; q\right) .$$

For these, by the procedure `qsumrecursion`, we get the recurrence equations

$$q^n(q^n-1)(q^n+q)p_n(x|q)+(q^{2n}-q)(q^{2n}x+xq^n+q^{n+1}x-2q^n+qx)p_{n-1}(x|q) \\ + \; q^n(q^n+1)(q^n-q)p_{n-2}(x|q) = 0 ,$$

and

$$q(q^n-1)(cq^n-1)(q^n+q)P_n(x;c;q) \\ + \; (q^{2n}-q)(q^{2n}x - 2q^{n+1} + q^{n+1}x - 2q^{m+1}c + xq^n + qx)P_{n-1}(x;c;q) \\ - q^n(q^n+1)(q^b-q)(q^n-cq)P_{n-2}(x;c;q) = 0 .$$

Exercise 21. Prove the identity

$${}_1\varphi_0\left(\begin{matrix} a \\ - \end{matrix}\,\middle|\, q; x\right) \cdot {}_1\varphi_0\left(\begin{matrix} b \\ - \end{matrix}\,\middle|\, q; ax\right) = {}_1\varphi_0\left(\begin{matrix} ab \\ - \end{matrix}\,\middle|\, q; x\right) .$$

Compute $p_n(1|q)$ in closed form.

With the algorithms of the **qsum** package some of the exercises of Dennis Stanton's contribution in these lecture notes can be solved.

Using Hahn's q-difference operator

$$D_q f(x) := \frac{f(x) - f(qx)}{(1-q)x} ,$$

one can also compute *q-difference equations* w.r.t. the variable x by the `qsumdiffeq` procedure.

5.3 Orthogonal polynomial solutions of recurrence equations

The classical orthogonal polynomials

$$P_n(x) = k_n\, x^n + \cdots$$

(Jacobi, Laguerre, Hermite and Bessel) satisfy a second order differential equation

$$\sigma(x)\, P_n''(x) + \tau(x)\, P_n'(x) + \lambda_n\, P_n(x) = 0 ,$$

where $\sigma(x)$ is a polynomial of degree at most 2 and $\tau(x)$ is a polynomial of degree 1.[3] From this differential equation one can determine the three-term recurrence equation for $P_n(x)$ in terms of the coefficients of $\sigma(x)$ and $\tau(x)$, see [12].

Using this information in the opposite direction, one can find the corresponding differential equation of second order—if applicable—from a given holonomic three-term recurrence equation. This is implemented in the procudure `REtoDE` of the **retode** package `retode.mpl` [13]. Note that Koornwinder and Swarttouw have a similar package **rec2ortho** [14] but use a different approach.

As an example, we consider the recurrence equation

$$P_{n+2}(x) - (x-n-1)\,P_{n+1}(x) + \alpha\,(n+1)^2\,P_n(x) = 0\;. \tag{5.1}$$

For this recurrence, the program finds that only if $\alpha = \frac{1}{4}$ there is a classical orthogonal polynomial solution $P_n(x)$ with $k_n = 1$ and density $\rho(x) = 4e^{-2x}$ on the interval $[-\frac{1}{2}, \infty)$ satisfying the differential equation

$$\left(x + \frac{1}{2}\right) P_n''(x) - 2x\,P_n'(x) + 2n\,P_n(x) = 0\;,$$

hence a translate of the Laguerre polynomials.

Similarly, the classical discrete orthogonal polynomials (Hahn, Meixner, Krawtchouk, Charlier) satisfy a second order difference equation

$$\sigma(x)\,\Delta\nabla P_n(x) + \tau(x)\Delta P_n(x) + \lambda_n\,P_n(x) = 0\;,$$

where $\nabla f(x) = f(x) - f(x-1)$ is the backward difference operator, $\sigma(x)$ is a polynomial of degree at most 2 and $\tau(x)$ is a polynomial of degree 1. Again, from this equation one can determine the three-term recurrence equation for $P_n(x)$ in terms of the coefficients of $\sigma(x)$ and $\tau(x)$ and convert this step, see [13]. This algorithm is accessible via `REtodiscreteDE`.

Taking again example (5.1), now we find that only for $\alpha < \frac{1}{4}$ there are classical discrete orthogonal polynomial solutions of the Meixner/Krawtchouk family.

[3] If $\sigma(x)$ has two different real roots $a < b$, then $P_n(x)$ is of the Jacobi family in the interval $[a, b]$, if it has one double root, then $P_n(x)$ is of the Bessel family, if it has one single root, $P_n(x)$ is of the Laguerre family, and if it is constant, then $P_n(x)$ is of the Hermite family.

Exercise 22. Find the classical orthogonal polynomial solutions of the recurrence equation

$$(n+3)\,P_{n+2}(x) - x(n+2)\,P_{n+1}(x) + (n+1)\,P_n(x) = 0\ .$$

Compute the recurrence equation for the functions

$$P_n(x) = {}_2F_0\left(\left.\begin{array}{c} -n, -x \\ - \end{array}\right| \lambda\right)$$

and determine whether they are classical orthogonal polynomial systems.

Finally, the classical q-orthogonal polynomials (see [8]) of the Hahn class satisfy a second order q-difference equation

$$\sigma(x)\,D_q D_{1/q} P_n(x) + \tau(x) D_q P_n(x) + \lambda_n\, P_n(x) = 0\ ,$$

where $\sigma(x)$ is a polynomial of degree at most 2 and $\tau(x)$ is a polynomial of degree 1. Again, from this equation one can determine the three-term recurrence equation for $P_n(x)$ in terms of the coefficients of $\sigma(x)$ and $\tau(x)$ and convert this step, see [13]. This algorithm is accessible via `REtoqDE`.

As an example, for the recurrence equation

$$P_{n+2}(x) - x\,P_{n+1}(x) + \alpha\, q^n\,(q^{n+1}-1)\,P_n(x) = 0\ ,$$

we get the corresponding q-difference equation

$$(x^2+\alpha)\,D_q D_{1/q} P_n(x) - \frac{x}{q-1} D_q P_n(x) + \frac{q(q^n-1)}{(q-1)^2\,q^n}\,P_n(x) = 0\ .$$

Exercise 23. Check that the little and big q-Legendre polynomials are in the Hahn class of q-orthogonal polynomials.

6 Epilogue

The author's Maple packages and their help pages can be downloaded from the website `http://www.mathematik.uni-kassel.de/~koepf/Publikationen`. Installation guidelines can be obtained by e-mail request. Finally, the accompanying Maple worksheets for this minicourse can be found at `http://www.mathematik.uni-kassel.de/~koepf/iivortrag.html`.

Software development is a time consuming activity! Software developers love it when their software is used. But they need also your support. Hence my suggestion: if you use one of the packages mentioned for your research, please cite its use!

References

1. Almkvist, G. and Zeilberger, D.: The method of differentiating under the integral sign. J. Symbolic Computation **10** (1990), 571–591.
2. Böing, H. and Koepf, W.: Algorithms for q-hypergeometric summation in computer algebra. J. Symbolic Computation **28** (1999), 777–799.
3. Bornemann, F.: private communication, 2002.
4. Char, B. W.: *Maple 8 Learning Guide*. Waterloo Maple, Waterloo, 2002.
5. Gasper, G. and Rahman, M.: *Basic Hypergeometric Series*. Encyclopedia of Mathematics and its Applications **35**, Cambridge University Press, London and New York, 1990.
6. Gosper Jr., R. W.: Decision procedure for indefinite hypergeometric summation. Proc. Natl. Acad. Sci. USA **75** (1978), 40–42.
7. Gruntz, D. and Koepf, W.: Maple package on formal power series. Maple Technical Newsletter **2** (2) (1995), 22–28.
8. Koekoek, R. and Swarttouw, R. F.: The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue. Report 98-17, Delft University of Technology, Faculty of Information Technology and Systems, Department of Technical Mathematics and Informatics, Delft; electronic version available at `http://aw.twi.tudelft.nl/~koekoek/askey`, 1998.
9. Koepf, W.: Power series in computer algebra. J. Symbolic Computation **13** (1992), 581–603.
10. Koepf, W.: Identities for families of orthogonal polynomials and special functions. Integral Transforms and Special Functions **5** (1997), 69–102.
11. Koepf, W.: *Hypergeometric Summation*. Vieweg, Braunschweig/Wiesbaden, 1998.
12. Koepf, W. and Schmersau, D.: Representations of orthogonal polynomials. J. Comput. Appl. Math. **90** (1998), 57–94.
13. Koepf, W. and Schmersau, D.: Recurrence equations and their classical orthogonal polynomial solutions. Appl. Math. Comput. **128** (2-3), special issue on Orthogonal Systems and Applications (2002), 303–327.
14. Koornwinder, T. H. and Swarttouw, R.: rec2ortho: an algorithm for identifying orthogonal polynomials given by their three-term recurrence relation as special functions. `http://turing.wins.uva.nl/~thk/recentpapers/rec2ortho.html`, 1996–1998
15. Le, H. Q., Abramov, S. A. and Geddes, K. O.: HypergeometricSum: A Maple package for finding closed forms of indefinite and definite sums of hypergeometric type. `ftp://cs-archive.uwaterloo.ca/cs-archive/CS-2001-24`, 2002.
16. Monagan, M. B. et al.: *Maple 8 Introductory Programming Guide*, Waterloo Maple, Waterloo, 2002.
17. Petkovšek, M., Wilf, H. and Zeilberger, D.: $A = B$. A. K. Peters, Wellesley, 1996.
18. Prudnikov, A. P., Brychkov, Yu. A. and Marichev, O. I.: *Integrals and Series. Vol. 3: More Special Functions*. Gordon & Breach, New York, 1990.
19. Salvy, B. and Zimmermann, P.: GFUN: A Maple package for the manipulation of generating and holonomic functions in one variable. ACM Transactions on Mathematical Software **20** (1994), 163–177.
20. Swarttouw, R.: CAOP: Computer algebra and orthogonal polynomials. `http://amstel.wins.uva.nl:7090/CAOP`, 1996–2002.

21. Van Hoeij, M.: Finite singularities and hypergeometric solutions of linear recurrence equations. J. Pure Appl. Algebra **139** (1999), 109–131.
22. Vidunas, R. and Koornwinder, T. H.: Zeilberger method for non-terminating hypergeometric series. 2002, to appear.
23. Wolfram's Research Mathematica Functions: `http://www.functions.wolfram.com`, 2002.
24. Zeilberger, D.: A fast algorithm for proving terminating hypergeometric identities. Discrete Math. **80** (1990), 207–211.

$3nj$-Coefficients and Orthogonal Polynomials of Hypergeometric Type

Joris Van der Jeugt

Department of Applied Mathematics and Computer Science, Ghent University, Krijgslaan 281-S9, B-9000 Gent, Belgium, `Joris.VanderJeugt@rug.ac.be`

Summary. We give a self-contained introduction to the theory of $3nj$-coefficients of $\mathfrak{su}(2)$ and $\mathfrak{su}(1,1)$, their hypergeometric expressions, and their relations to orthogonal polynomials. The $3nj$-coefficients of $\mathfrak{su}(2)$ play a crucial role in various physical applications (dealing with the quantization of angular momentum), but here we shall deal with their mathematical importance only.

We begin with a brief explanation of representations of the Lie algebra $\mathfrak{su}(2)$, and of the tensor product decomposition. All this is derived in a purely algebraic context (avoiding the notion of Lie groups for those who are not familiar with this). The Clebsch-Gordan coefficients (or $3j$-coefficients) are defined, and their expression as a hypergeometric series is deduced. The relation with (discrete) Hahn polynomials is emphasized.

The tensor product of three representations is discussed, and the relevant Racah coefficients (or $6j$-coefficients) are defined. The explicit expression of a Racah coefficient as a hypergeometric series of ${}_4F_3$-type is derived. The connection with Racah polynomials and their orthogonality is also determined.

As a second example closely related to $\mathfrak{su}(2)$, a class of representations (together with their $3j$- and $6j$-coefficients) of the Lie algebra $\mathfrak{su}(1,1)$ is considered.

Finally, we introduce general coupling theory : the tensor product of $(n+1)$ representations is considered, and "generalized recoupling coefficients" or $3nj$-coefficients are defined. In this context, the Biedenharn-Elliott identity is shown to play a crucial role. We shall derive various results from this identity, e.g. convolution theorems for certain orthogonal polynomials, orthogonal polynomials in several variables, and interpretations of $3nj$-coefficients as connection coefficients.

1 The Lie algebras $\mathfrak{su}(2)$ and $\mathfrak{su}(1,1)$ and representations 27

1.1 Lie algebras, enveloping algebras 27
1.2 Representations or modules 28
1.3 Verma modules of $\mathfrak{sl}(2,\mathbb{C})$ 29
1.4 $\star$-operations and $\star$-representations 30
1.5 $\star$-representations D_j of $\mathfrak{su}(2)$ 31
1.6 $\star$-representations of $\mathfrak{su}(1,1)$ 33
1.7 Notes and Exercises 34

E. Koelink and W. Van Assche (Eds.): LNM 1817, pp. 25–92, 2003.

2 Hypergeometric series and some transformation formulas . 35

2.1 Hypergeometric series; Vandermonde's theorem . . . 35
2.2 Generalized hypergeometric series . . . 37
2.3 Notes and Exercises . . . 39

3 Tensor product of two $\mathfrak{su}(2)$ representations . . . 40

3.1 A realization of $\mathfrak{su}(2)$ . . . 40
3.2 Tensor product decomposition . . . 41
3.3 Some expressions for Clebsch-Gordan coefficients of $\mathfrak{su}(2)$ . . . 45
3.4 Symmetries of $\mathfrak{su}(2)$ Clebsch-Gordan coefficients . . . 47
3.5 Orthogonality relations . . . 48
3.6 Recurrence relations for Clebsch-Gordan coefficients . . . 49
3.7 Hahn and dual Hahn polynomials . . . 49
3.8 Recurrence relations . . . 51
3.9 Tensor product of $\mathfrak{su}(1,1)$ representations . . . 52
3.10 Notes and Exercises . . . 54

4 Tensor product of three $\mathfrak{su}(2)$ representations and Racah coefficients . . . 56

4.1 Definition of the Racah coefficient . . . 56
4.2 The $6j$-coefficient . . . 59
4.3 Expressions for the Racah coefficient . . . 60
4.4 The Racah polynomial . . . 64
4.5 Wilson polynomials . . . 66
4.6 Racah coefficients of $\mathfrak{su}(1,1)$ . . . 68
4.7 Notes and Exercises . . . 69

5 Tensor product of $n+1$ representations and $3nj$-coefficients 70

5.1 $3nj$-coefficients . . . 70
5.2 An example : the Biedenharn-Elliott identity . . . 73
5.3 A recurrence relation for Racah polynomials . . . 76
5.4 Expressions for $3nj$-coefficients . . . 77
5.5 Multivariable Jacobi polynomials associated with binary coupling trees . . . 81
5.6 Notes and Exercises . . . 88

References . . . 89

1 The Lie algebras $\mathfrak{su}(2)$ and $\mathfrak{su}(1,1)$ and representations

1.1 Lie algebras, enveloping algebras

Although we shall be concerned only with the Lie algebras $\mathfrak{su}(2)$ and $\mathfrak{su}(1,1)$, let us first consider the general definition of a Lie algebra $\mathfrak{g}$. The ground field K is $\mathbb{R}$ or $\mathbb{C}$.

Definition 1.1. *A vector space* $\mathfrak{g}$ *over the field* K, *with a bilinear operation* $[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, *is called a Lie algebra over* K *if*

- $[x, y] = -[y, x]$ *(anti-symmetry)*
- $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ *(Jacobi-identity)*

for $x, y, z \in \mathfrak{g}$.

It is obvious how to define a *subalgebra* $\mathfrak{h}$ of $\mathfrak{g}$: a subspace $\mathfrak{h}$ of $\mathfrak{g}$ is a subalgebra if $[x, y] \in \mathfrak{h}$ whenever $x, y \in \mathfrak{h}$.

The operation $[\ ,\ ]$ is usually referred to as the *bracket* or *commutator*. The reason for the last terminology comes from the following observation. Let $\mathfrak{g}$ be an ordinary (associative) algebra over K, with the product of two elements $x, y \in \mathfrak{g}$ denoted by $x \cdot y \equiv xy$. Then $\mathfrak{g}$ is turned into a Lie algebra by taking

$$[x, y] = xy - yx. \tag{1.1}$$

The simplest example of a Lie algebra is $\mathfrak{gl}(n, \mathbb{C})$. As a vector space, $\mathfrak{gl}(n, \mathbb{C})$ consists of all complex $(n \times n)$ matrices. The product is the usual matrix product; and the bracket is defined by (1.1). As a basis for $\mathfrak{gl}(n, \mathbb{C})$, one can take the n^2 matrix units e_{ij} $(i, j = 1, \ldots, n)$. The commutator for these basis elements reads :

$$[e_{ij}, e_{kl}] = \delta_{jk} e_{il} - \delta_{il} e_{kj}.$$

Since the trace of square matrices satisfies $\mathrm{tr}(xy) = \mathrm{tr}(yx)$, we have that $\mathrm{tr}([x, y]) = 0$ for all elements of $\mathfrak{gl}(n, \mathbb{C})$. Thus, the subspace consisting of traceless elements of $\mathfrak{gl}(n, \mathbb{C})$ forms a Lie subalgebra, denoted by $\mathfrak{sl}(n, \mathbb{C})$. Clearly, $\mathfrak{sl}(n, \mathbb{C})$ has $n^2 - 1$ basis elements.

More generally, let V be a vector space over K, and denote by $\mathrm{End}(V)$ the set of linear transformations from V to V. $\mathrm{End}(V)$ is a ring with the usual product operation. Defining the bracket by (1.1), $\mathrm{End}(V)$ becomes a Lie algebra over K. Usually, this Lie algebra will be denoted by $\mathfrak{gl}(V)$.

Let us consider the Lie algebra $\mathfrak{sl}(2, \mathbb{C})$, with the following standard basis :

$$J_0 = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}, \qquad J_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad J_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{1.2}$$

The basic commutation relations then read :

$$[J_0, J_\pm] = \pm J_\pm, \qquad [J_+, J_-] = 2J_0. \tag{1.3}$$

Let $\mathfrak{g}$ be a Lie algebra, say with basis elements x_i $(i = 1, 2 \ldots, m)$. The *structure constants* c_{ij}^k are the constants appearing in the commutators

$$[x_i, x_j] = \sum_k c_{ij}^k x_k.$$

The *enveloping algebra* $\mathcal{U}(\mathfrak{g})$ is the associative algebra with unit, generated by the elements x_i, and subject to the relations

$$x_i x_j - x_j x_i - \sum_k c_{ij}^k x_k = 0.$$

The structure of $\mathcal{U}(\mathfrak{g})$ is independent of a basis (it is universal), but we shall not be concerned with the details of this. For our further developments, we only need some properties of $\mathcal{U}(\mathfrak{sl}(2, \mathbb{C}))$.

Proposition 1.1. *The following relations hold in* $\mathcal{U}(\mathfrak{sl}(2, \mathbb{C}))$ *(*$n \in \mathbb{N}$*) :*

(a) $J_0 J_\pm^n = J_\pm^n (J_0 \pm n)$
(b) $[J_-, J_+^n] = -n(n-1) J_+^{n-1} - 2n J_+^{n-1} J_0$
(c) $[C, x] = 0$ *for all* $x \in \mathcal{U}(\mathfrak{sl}(2, \mathbb{C}))$*, where*

$$C = J_+ J_- + J_0^2 - J_0 = J_- J_+ + J_0^2 + J_0$$

(the Casimir operator).

Proof. For $n = 1$, (a) follows directly from $[J_0, J_\pm] = \pm J_\pm$; next use induction on n. Similarly for (b), the case $n = 1$ is trivial. Then, use the general commutator identity $[x, yz] = y[x, z] + [x, y]z$ to write

$$[J_-, J_+^n] = [J_-, J_+^{n-1} J_+] = J_+^{n-1}[J_-, J_+] + [J_-, J_+^{n-1}] J_+,$$

and then use again induction on n. To prove (c), it is sufficient to show that it holds for x equal to J_0, J_+ and J_-. Each of these is easy to verify, e.g.

$$\begin{aligned}[C, J_+] &= [J_+ J_- + J_0^2 - J_0, J_+] \\ &= J_+[J_-, J_+] + J_0[J_0, J_+] + [J_0, J_+]J_0 - [J_0, J_+] \\ &= -2J_+ J_0 + J_0 J_+ + J_+ J_0 - J_+ = 0.\end{aligned}$$

1.2 Representations or modules

Let us now come to the topic of representations of a Lie algebra $\mathfrak{g}$, or $\mathfrak{g}$-modules. First, a linear transformation $\varphi : \mathfrak{g}_1 \to \mathfrak{g}_2$ ($\mathfrak{g}_1$ and $\mathfrak{g}_2$ two Lie algebras over K) is called a *homomorphism* if $\varphi([x, y]) = [\varphi(x), \varphi(y)]$ for all $x, y \in \mathfrak{g}_1$. A *representation* of a Lie algebra $\mathfrak{g}$ in V (with V a vector space over K) is a homomorphism $\varphi : \mathfrak{g} \to \mathfrak{gl}(V)$. V is then called the representation space. It is often convenient to use the language of modules along with the equivalent language of representations.

Definition 1.2. *A vector space* V *endowed with an operation* $\mathfrak{g} \times V \to V$ *(denoted by* $(x, v) \to x{\cdot}v \equiv xv$*) is called a* $\mathfrak{g}$-module *if the following conditions are satisfied (*$x, y \in \mathfrak{g}; v, w \in V; a, b \in K$*) :*

(m1) $(ax + by) \cdot v = a(x \cdot v) + b(y \cdot v)$,
(m2) $x \cdot (av + bw) = a(x \cdot v) + b(x \cdot w)$,
(m3) $[x, y] \cdot v = x \cdot y \cdot v - y \cdot x \cdot v$.

The relation between representations and $\mathfrak{g}$-modules is natural : if φ is a representation of $\mathfrak{g}$ in the vector space V, then V is a $\mathfrak{g}$-module under the action $x \cdot v = \varphi(x)(v)$ (and vice versa). Also the notions of $\mathfrak{g}$-module and $\mathcal{U}(\mathfrak{g})$-module are equivalent.

If W is a subspace of a $\mathfrak{g}$-module V such that $xw \in W$ for every $x \in \mathfrak{g}$ and $w \in W$, then W is a $\mathfrak{g}$-submodule. A $\mathfrak{g}$-module V which has only the trivial $\mathfrak{g}$-submodules V and $\{0\}$ is called *irreducible.* Often, this is referred to as a *simple* $\mathfrak{g}$-module V. V is called *completely reducible* if V is a direct sum of irreducible $\mathfrak{g}$-submodules. In other words, V is completely reducible if each $\mathfrak{g}$-submodule W of V has a complement W' which is itself a $\mathfrak{g}$-submodule with $V = W \oplus W'$. All this terminology (irreducible, completely reducible) applies to representations as well.

An important notion to introduce is the concept of *tensor product* of $\mathfrak{g}$-modules. Let V and W be $\mathfrak{g}$-modules, and let $V \otimes W$ be the tensor product over K of the underlying vector spaces. Recall that if V and W have respective bases $v_1, v_2, \ldots$ and $w_1, w_2, \ldots$, then $V \otimes W$ has a basis consisting of the vectors $v_i \otimes w_j$. Now $V \otimes W$ has the structure of a $\mathfrak{g}$-module by defining :

$$x \cdot (v \otimes w) = x \cdot v \otimes w + v \otimes x \cdot w. \tag{1.4}$$

To convince yourself that this is the right definition, the crucial condition to verify is (m3).

1.3 Verma modules of $\mathfrak{sl}(2, \mathbb{C})$

Now we shall concentrate on some special representations (or modules) of $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{C})$. A *Verma module* is a $\mathfrak{g}$-module generated by a single vector v. In particular, let v be a vector satisfying

$$J_0 v = \lambda v,$$
$$J_- v = 0.$$

We say that the action of J_0 on v is diagonal, and that v is annihilated by J_-. Let us consider the Verma module V generated by the single vector v. Clearly, the vectors $v, J_+v, J_+^2 v, \ldots$ are vectors of V. Let us compute the action of J_0, J_+ and J_- on such a vector $J_+^n v$. Using Proposition 1.1, we find :

$$\begin{aligned}
J_0(J_+^n v) &= J_+^n(J_0+n)v = (\lambda+n)(J_+^n v),\\
J_+(J_+^n v) &= J_+^{n+1}v,\\
J_-(J_+^n v) &= (J_+^n J_- - n(n-1)J_+^{n-1} - 2nJ_+^{n-1}J_0)v\\
&= -n(2\lambda+n-1)(J_+^{n-1}v) \qquad (n\geq 1).
\end{aligned}$$

Thus it is clear that the vectors $J_+^n v$ $(n\in\mathbb{N})$ form a basis of the Verma module V. The eigenvalue of $J_+^n v$ under the action of J_0 is $\lambda+n$. Let us denote these vectors by

$$v_{\lambda+n} = J_+^n v.$$

Then the previous computations show :

Proposition 1.2. *The $\mathfrak{sl}(2,\mathbb{C})$ Verma module $V=V_\lambda$ generated by a single vector $v=v_\lambda$ satisfying $J_0 v=\lambda v$ and $J_- v = 0$ has a basis of vectors $v_{\lambda+n}=J_+^n v$ $(n\in\mathbb{N})$, and the action is given by*

$$\begin{aligned}
J_0 v_{\lambda+n} &= (\lambda+n)v_{\lambda+n},\\
J_+ v_{\lambda+n} &= v_{\lambda+n+1},\\
J_- v_{\lambda+n} &= -n(2\lambda+n-1)v_{\lambda+n-1}.
\end{aligned}$$

Observe that

$$Cv_{\lambda+n} = CJ_+^n v_\lambda = J_+^n C v_\lambda = (\lambda^2-\lambda)v_{\lambda+n}.$$

From the actions given in this proposition, it is clear that V_λ is a simple (or irreducible) $\mathfrak{sl}(2,\mathbb{C})$-module if $2\lambda-1$ is different from $-1,-2,-3,\ldots$. So, if $-2\lambda\notin\mathbb{N}$, then V_λ is simple. If $-2\lambda\in\mathbb{N}$, then V_λ is not simple. Let us consider this case in more detail. Suppose $\lambda=-j$ with $j\in\frac{1}{2}\mathbb{N}$. Then $J_- v_{j+1}=0$, so the subspace M_λ of V_λ consisting of vectors v_{j+n+1} $(n\in\mathbb{N})$ is a $\mathfrak{sl}(2,\mathbb{C})$-submodule of V_λ. The *quotient module* $L_\lambda \equiv V_\lambda/M_\lambda$ is then finite-dimensional, and is spanned by (representatives of) the $(2j+1)$ vectors v_m with $m=-j,-j+1,\ldots,j$. This module is often denoted by D_j, with $j\in\frac{1}{2}\mathbb{N}$. Thus, the action of $\mathfrak{sl}(2,\mathbb{C})$ on the basis of D_j is given by

$$\begin{aligned}
&J_0 v_m = m v_m, \qquad (m=-j,-j+1,\ldots,j)\\
&J_+ v_m = v_{m+1}, \qquad (m=-j,-j+1,\ldots,j-1), \qquad J_+ v_j = 0,\\
&J_- v_m = (j+m)(j-m+1)v_{m-1} \qquad (m=-j,-j+1,\ldots,j).
\end{aligned}$$

1.4 $\star$-operations and $\star$-representations

In order to introduce an inner product on the representation space, we need the notions of $\star$-algebra and $\star$-representation.

Definition 1.3. *The Lie algebra $\mathfrak{g}$ over K is a $\star$-algebra if there exists a conjugate-linear anti-automorphic involution $\star : \mathfrak{g}\to\mathfrak{g} : x\to x^*$, satisfying*

$$\begin{aligned}(x^*)^* &= x\\ (ax+by)^* &= \bar{a}x^* + \bar{b}y^*\\ [x,y]^* &= [y^*, x^*]\end{aligned}$$

with $x, y \in \mathfrak{g}$ *and* $a, b \in K$.

For $\mathfrak{sl}(2,\mathbb{C})$, there exist two non-equivalent $\star$-operations. These are explicitly given by

$$J_0^* = J_0, \qquad J_\pm^* = J_\mp \qquad (\mathfrak{su}(2) \text{ case}) \tag{1.5}$$

and

$$J_0^* = J_0, \qquad J_\pm^* = -J_\mp \qquad (\mathfrak{su}(1,1) \text{ case}). \tag{1.6}$$

The *real* Lie algebra with commutators (1.3) and $\star$-operation (1.5) is called the Lie algebra $\mathfrak{su}(2)$; the *real* Lie algebra with commutators (1.3) and $\star$-operation (1.6) is called the Lie algebra $\mathfrak{su}(1,1)$.

Suppose a $\star$-operation is given on a Lie algebra $\mathfrak{g}$. Consider a $\mathfrak{g}$-module V (a representation), and suppose a Hermitian form $\langle\,,\,\rangle$ is given on V, i.e. $\langle\,,\,\rangle : V \times V \to \mathbb{C}$ such that

$$\langle u, v\rangle = \overline{\langle v, u\rangle}, \text{ and } \langle u, av+bw\rangle = a\langle u, v\rangle + b\langle u, w\rangle \qquad (a, b \in \mathbb{C}; u, v, w \in V);$$

in the real case, the Hermitian form is just a bilinear form. Then V is a $\star$*-representation* if for every $x \in \mathfrak{g}$ and all $v, w \in V$,

$$\langle x \cdot v, w\rangle = \langle v, x^* \cdot w\rangle.$$

Let us now return to the representations of $\mathfrak{sl}(2,\mathbb{C})$, and investigate when these representations are $\star$-representations for $\mathfrak{su}(2)$ or $\mathfrak{su}(1,1)$. In particular, we will be interested in the case that the Hermitian form on such a representation V coincides with an inner product. For this, it is sufficient that $\langle v, v\rangle > 0$ for every basis vector of V.

1.5 $\star$-representations D_j of $\mathfrak{su}(2)$

Let us first consider the $\star$-operation associated with $\mathfrak{su}(2)$. Consider the Verma module V_λ (or its submodule, or the quotient module). On V_λ, one can define a Hermitian form by requiring

$$\langle v_\lambda, v_\lambda\rangle = 1,$$

and postulating that

$$\langle xv, w\rangle = \langle v, x^* w\rangle$$

for all $v, w \in V_\lambda$ and $x \in \mathfrak{su}(2)$. From $\langle J_0 v_\lambda, v_\lambda\rangle = \langle v_\lambda, J_0 v_\lambda\rangle$ it follows that λ has to be real. Furthermore,

$$\begin{aligned}\langle v_{\lambda+n}, v_{\lambda+n}\rangle &= \langle J_+ v_{\lambda+n-1}, v_{\lambda+n}\rangle = \langle v_{\lambda+n-1}, J_- v_{\lambda+n}\rangle\\ &= -n(2\lambda+n-1)\langle v_{\lambda+n-1}, v_{\lambda+n-1}\rangle,\end{aligned}$$

or

$$\langle v_{\lambda+n}, v_{\lambda+n}\rangle = n!(-2\lambda)(-2\lambda-1)\cdots(-2\lambda-n+1).$$

So if λ is positive, the sign of $\langle v_{\lambda+n}, v_{\lambda+n}\rangle$ changes according to even or odd n-values, and the Hermitian form on the Verma module is not an inner product.

Consider the case that λ is negative. Then $\langle v_{\lambda+n}, v_{\lambda+n}\rangle > 0$ if $n < 1-2\lambda$ and has alternating signs for $n > 1-2\lambda$. So, again the Hermitian form on the Verma module is not an inner product. However, in the case that $\lambda = -j$ with $j \in \frac{1}{2}\mathbb{N}$, the form does yield an inner product on the quotient module $L_\lambda = V_\lambda/M_\lambda \equiv D_j$. Then we can write

$$\begin{aligned}\langle v_m, v_m\rangle &= (j+m)!(2j)(2j-1)\cdots(j-m+1)\langle v_{-j}, v_{-j}\rangle \\ &= \frac{(j+m)!(2j)!}{(j-m)!},\end{aligned}$$

for $m = -j, -j+1, \ldots, j$. So we can define an *orthonormal basis* for the representation D_j by putting

$$e_m \equiv e_m^{(j)} = \sqrt{\frac{(j-m)!}{(j+m)!(2j)!}}\, v_m.$$

We come to the following conclusion :

Theorem 1.1. *For $j \in \frac{1}{2}\mathbb{N}$, D_j is an irreducible $\star$-representation of $\mathfrak{su}(2)$ of dimension $2j+1$, with inner product $\langle e_m^{(j)}, e_{m'}^{(j)}\rangle = \delta_{m,m'}$ for the basis $e_m^{(j)}$ ($m = -j, -j+1, \ldots, j$), and with action given by :*

$$\begin{aligned} J_0 e_m^{(j)} &= m e_m^{(j)}, \\ J_+ e_m^{(j)} &= \sqrt{(j-m)(j+m+1)}\, e_{m+1}^{(j)}, \\ J_- e_m^{(j)} &= \sqrt{(j+m)(j-m+1)}\, e_{m-1}^{(j)}. \end{aligned} \tag{1.7}$$

For the Casimir operator, one has $C^* = C$ and $Ce_m^{(j)} = j(j+1)e_m^{(j)}$.

The m-value of the vector $e_m^{(j)}$ is often referred to as the *weight* of the vector. All vectors can be obtained by acting on the *highest weight vector* $e_j^{(j)}$ with powers of J_-, or by acting on the *lowest weight vector* $e_{-j}^{(j)}$ with powers of J_+; e.g.

$$e_m^{(j)} = \sqrt{\frac{(j+m)!}{(2j)!(j-m)!}}(J_-)^{j-m} e_j^{(j)} \tag{1.8}$$

$$= \sqrt{\frac{(j-m)!}{(2j)!(j+m)!}}(J_+)^{j+m} e_{-j}^{(j)}. \tag{1.9}$$

1.6 $\star$-representations of $\mathfrak{su}(1,1)$

Secondly, consider the $\star$-operation associated with $\mathfrak{su}(1,1)$. Consider again the Verma module V_λ with Hermitian form defined by

$$\langle v_\lambda, v_\lambda \rangle = 1,$$

and

$$\langle xv, w \rangle = \langle v, x^* w \rangle$$

for all $v, w, \in V_\lambda$ and $x \in \mathfrak{su}(1,1)$. Again, λ has to be real. Now we get :

$$\begin{aligned} \langle v_{\lambda+n}, v_{\lambda+n} \rangle &= \langle J_+ v_{\lambda+n-1}, v_{\lambda+n} \rangle = -\langle v_{\lambda+n-1}, J_- v_{\lambda+n} \rangle \\ &= n(2\lambda+n-1)\langle v_{\lambda+n-1}, v_{\lambda+n-1} \rangle, \end{aligned}$$

or

$$\langle v_{\lambda+n}, v_{\lambda+n} \rangle = n!(2\lambda)(2\lambda+1)\cdots(2\lambda+n-1).$$

If λ is negative, the sign changes again and the Hermitian form is not an inner product. If $\lambda > 0$, the sign of $\langle v_{\lambda+n}, v_{\lambda+n} \rangle$ is always positive, and the Hermitian form is an inner product. Recall that in that case the Verma module V_λ is also irreducible. Using the common notation for the *Pochhammer symbol*

$$(\alpha)_n = \alpha(\alpha+1)\cdots(\alpha+n-1) \qquad (n \in \mathbb{N}_+), \qquad (a)_0 = 1, \tag{1.10}$$

we have

$$\langle v_{\lambda+n}, v_{\lambda+n} \rangle = n!(2\lambda)_n.$$

Define an orthonormal basis

$$e_n^{(\lambda)} \equiv e_n = \frac{v_{\lambda+n}}{\sqrt{n!(2\lambda)_n}}.$$

Now we have :

Theorem 1.2. *For $\lambda > 0$, V_λ is an irreducible $\star$-representation of $\mathfrak{su}(1,1)$, with inner product $\langle e_n^{(\lambda)}, e_{n'}^{(\lambda)} \rangle = \delta_{n,n'}$ for the basis $e_n^{(\lambda)}$ ($n \in \mathbb{N}$), and with action given by :*

$$\begin{aligned} J_0 e_n^{(\lambda)} &= (\lambda+n) e_n^{(\lambda)}, \\ J_+ e_n^{(\lambda)} &= \sqrt{(n+1)(2\lambda+n)}\, e_{n+1}^{(\lambda)}, \\ J_- e_n^{(\lambda)} &= -\sqrt{n(2\lambda+n-1)}\, e_{n-1}^{(\lambda)}. \end{aligned} \tag{1.11}$$

The Casimir operator satisfies $C^* = C$ and $C e_n^{(\lambda)} = \lambda(\lambda-1) e_n^{(\lambda)}$.

1.7 Notes and Exercises

Lie groups and Lie algebras play an important role in mathematics and mathematical physics. Some classical references on Lie algebras and/or their representations are [9], [19], [20], [42], [49]. Textbooks on Lie groups or Lie algebras are mainly devoted to the classification of simple Lie groups and algebras. The subject arose out of the work of Sophus Lie on differential equations and contact transformations. The algebraic problems in the classification theory were solved by W. Killing and E. Cartan around 1889-1913; then came further contributions by H. Weyl, E. Cartan, E.B. Dynkin, A. Weil, C. Chevalley, A. Borel, ...

Lie groups and algebras are of fundamental importance in physics, especially since the development of quantum mechanics. The first applications of continuous groups come from representations of the Lorentz group and the Poincaré group. Gradually, Lie groups made their way to nuclear physics, elementary particle physics, quantum field theory, ...

In this section, we have considered only the Lie algebras $\mathfrak{su}(2)$ and $\mathfrak{su}(1,1)$. The Lie algebra $\mathfrak{su}(2)$ and its representations is one of the basic examples. It also coincides with the angular momentum algebra in quantum physics, which is why physicists have paid so much attention to it. Much of the developments of $\mathfrak{su}(2)$ representation theory was performed by physicists [32], [52], [33], [41], [5], see also [8].

It is uncommon to treat $\mathfrak{su}(2)$ representations by means of Verma modules, a technique appropriate for the analysis of highest weight modules of Lie algebras or Kac-Moody algebras in general [22]. Here, we have chosen for this approach because it allows us to analyse representations of $\mathfrak{su}(2)$ and $\mathfrak{su}(1,1)$ at the same time.

Exercises

1. In section 1.1 the Lie algebras $\mathfrak{gl}(n,\mathbb{R})$ and $\mathfrak{sl}(n,\mathbb{R})$ were introduced. Let $V=\mathbb{R}^{2n}$, and define a nondegenerate skew-symmetric form f on V by the matrix
$$B=\begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$
(even-dimensionality is a necessary condition for the existence of a nondegenerate bilinear form satisfying $f(v,w)=-f(w,v)$ for all $v,w,\in V$). Denote by $\mathfrak{sp}(2n)$ the set of all endomorphisms of V satisfying $f(x(v),w)=-f(v,x(w))$. In matrix notation, $\mathfrak{sp}(2n)$ consists of $(2n\times 2n)$-matrices X satisfying $X^tB+BX=0$. Show that $\mathfrak{sp}(2n)$ is a Lie subalgebra of $\mathfrak{sl}(2n,\mathbb{R})$; it is called the *symplectic Lie algebra*. Show that the dimension of $\mathfrak{sp}(2n)$ is $2n^2+n$ (e.g. by constructing a basis).
2. Let $V=\mathbb{R}^{2n+1}$, and define a nondegenerate symmetric form f on V by the matrix

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_n \\ 0 & I_n & 0 \end{pmatrix}.$$

Denote by $\mathfrak{so}(2n+1)$ the set of all endomorphisms of V satisfying $f(x(v),w) = -f(v,x(w))$. In matrix notation, $\mathfrak{so}(2n+1)$ consists of matrices X satisfying $X^tB + BX = 0$. Show that $\mathfrak{so}(2n+1)$ is a Lie subalgebra of $\mathfrak{sl}(2n+1,\mathbb{R})$; it is called the *orthogonal Lie algebra*. Show that the dimension of $\mathfrak{so}(2n+1)$ is also $2n^2+n$.

3. Let again $V = \mathbb{R}^{2n}$, and define a nondegenerate symmetric form f on V by the matrix

$$B = \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix}.$$

The set of all endomorphisms of V satisfying $f(x(v),w) = -f(v,x(w))$ is denoted by $\mathfrak{so}(2n)$. Show that $\mathfrak{so}(2n)$ is a Lie subalgebra of $\mathfrak{sl}(2n,\mathbb{R})$ (also an orthogonal Lie algebra) of dimension $2n^2-n$.

4. Show, as in Proposition 1.1, that

$$[J_+, J_-^n] = -n(n-1)J_-^{n-1} + 2nJ_-^{n-1}J_0.$$

5. Verify that the action on $V \otimes W$, as defined in (1.4), turns $V \otimes W$ into a $\mathfrak{g}$-module (by checking (m1)–(m3) of definition 1.2).
6. Let $\mathfrak{g}$ be a Lie algebra. Show that ad is a representation of $\mathfrak{g}$ into itself (the *adjoint representation*) by

$$\mathrm{ad}(x)\cdot y = [x,y], \qquad (x,y \in \mathfrak{g}).$$

7. In section 1.3 the Verma modules of $\mathfrak{sl}(2,\mathbb{C})$ with lowest weight λ were constructed. Use the same technique to construct the Verma modules W_λ with highest weight λ. In this case, the generating vector v satisfies $J_0v = \lambda v$ and $J_+v = 0$. Construct a basis for W_λ, and determine the submodule (if there exists one).

2 Hypergeometric series and some transformation formulas

2.1 Hypergeometric series; Vandermonde's theorem

Recall the definition of the hypergeometric series

$$F(a,b;c;z) = {}_2F_1(a,b;c;z) = {}_2F_1\left({a,\, b \atop c}; z\right) = \sum_{k=0}^{\infty} \frac{(a)_k(b)_k}{k!(c)_k} z^k,$$

where $(a)_k$ is the Pochhammer symbol (1.10); sometimes we write $(a,b,\ldots)_k$ for $(a)_k(b)_k\ldots$. The convergence properties for such series are well known; for

$|z| < 1$, the series converges absolutely. Here, we shall mainly be dealing with *terminating series* : if one of the numerator parameters a or b is a negative integer $-n$, then the series consists of $n+1$ terms and terminates. In general, the denominator parameter c is not a negative integer. In case of a terminating series, c is allowed to be a negative integer which must be smaller than the negative integer in the numerator responsible for the termination.

Hypergeometric series appear as solutions of a second order differential equation of the form

$$z(1-z)\frac{d^2y}{dz^2} + [c-(a+b+1)z]\frac{dy}{dz} - aby = 0. \tag{2.1}$$

The most general solution for $|z| < 1$ is given by

$$y = A\ F(a,b;c;z) + Bz^{1-c}\ F(a+1-c, b+1-c; 2-c; z). \tag{2.2}$$

One can verify that also $(1-z)^{c-a-b}F(c-a,c-b;c;z)$ satisfies (2.1), hence it can be expressed in terms of (2.2). By comparing coefficients, one finds :

$$F(a,b;c;z) = (1-z)^{c-a-b}F(c-a,c-b;c;z); \tag{2.3}$$

this is known as *Euler's transformation formula.*

Let u and v be two positive integers. All combinations of n elements from the set of $u+v$ elements can be obtained by taking k elements from the first set of u elements and combining these with $n-k$ elements from the second set of v elements. Thus :

$$\sum_{k=0}^{n} \binom{u}{k}\binom{v}{n-k} = \binom{u+v}{n}. \tag{2.4}$$

This is now a polynomial expression in u and v, so it is valid in general. One can rewrite the binomial coefficients by means of

$$\binom{u}{k} = \frac{u(u-1)\dots(u-k+1)}{k!} = (-1)^k\frac{(-u)_k}{k!},$$

and then (2.4) becomes

$$\sum_{k=0}^{n} \binom{n}{k}(-u)_k(-v)_{n-k} = (-u-v)_n. \tag{2.5}$$

Using trivial Pochhammer identities such as

$$(-v)_{n-k} = \frac{(-1)^k(-v)_n}{(1+v-n)_k},$$

(2.5) becomes

$$\sum_{k=0}^{n} \frac{(-n)_k(-u)_k(-v)_n}{k!(1+v-n)_k} = (-u-v)_n,$$

or

$${}_2F_1\left({-n,-u \atop 1+v-n};1\right) = \frac{(-u-v)_n}{(-v)_n};$$

this can be rewritten as

Theorem 2.1. *Let n be a nonnegative integer, then*

$${}_2F_1\left({-n,b \atop c};1\right) = \frac{(c-b)_n}{(c)_n}. \tag{2.6}$$

This is known as Vandermonde's summation formula for a terminating ${}_2F_1$ of unit argument.

2.2 Generalized hypergeometric series

The ordinary hypergeometric series ${}_2F_1$ has two numerator parameters a and b, and one denominator parameter c. More generally, one can consider the *generalized hypergeometric series*

$${}_pF_q\left({a_1,\ldots,a_p \atop b_1,\ldots,b_q};z\right) = \sum_{k=0}^{\infty} \frac{(a_1)_k\cdots(a_p)_k}{k!(b_1)_k\cdots(b_q)_k} z^k. \tag{2.7}$$

In the for us interesting case when $p=q+1$, the series converges when $|z|<1$. Most of the series that we will consider here will be terminating anyway.

For the generalized hypergeometric series ${}_3F_2$ of unit argument, a summation formula follows from Euler's formula (2.3). Indeed, consider the expansion of the right hand side of (2.3) as a series in z :

$$(1-z)^{c-a-b}F(c-a,c-b;c;z) = \sum_k \binom{c-a-b}{k}(-1)^k z^k \sum_l \frac{(c-a)_l(c-b)_l}{l!(c)_l} z^l.$$

The coefficient of z^n in this expression is

$$\sum_l \frac{(c-a)_l(c-b)_l(a+b-c)_{n-l}}{l!(c)_l(n-l)!}.$$

Comparing this with the coefficient of z^n in the left hand side of (2.3), gives

$$\sum_l \frac{(c-a)_l(c-b)_l(-n)_l(a+b-c)_n}{l!(c)_l(1-a-b+c-n)_l n!} = \frac{(a)_n(b)_n}{(c)_n n!},$$

or :

$${}_3F_2\left({c-a,c-b,-n \atop c,1-a-b+c-n};1\right) = \frac{(a)_n(b)_n}{(c)_n(a+b-c)_n}. \tag{2.8}$$

In other words :

Theorem 2.2. *Let n be a nonnegative integer, and $a+b-n+1=c+d$, then*

$$ {}_3F_2\left({a,b,-n \atop c,d};1\right)=\frac{(c-a)_n(c-b)_n}{(c)_n(c-a-b)_n}. \tag{2.9}$$

This is known as the *Pfaff-Saalschütz summation formula.*

Next, we deduce a transformation formula for a terminating balanced ${}_4F_3$ series of unit argument, due to Whipple (we follow the proof of [1]).

Theorem 2.3.

$$\begin{aligned} &{}_4F_3\left({-n,a,b,c \atop d,e,f};1\right)= \\ &\qquad\frac{(e-c)_n(f-c)_n}{(e)_n(f)_n}\;{}_4F_3\left({-n,d-a,d-b,c \atop d,d+e-a-b,d+f-a-b};1\right), \end{aligned} \tag{2.10}$$

where $a+b+c-n+1=d+e+f$ (balance condition).

Proof. Let a,b,c,d,e,f be arbitrary parameters. The coefficient of z^n in

$$ {}_2F_1\left({a,b \atop c};z\right){}_2F_1\left({d,e \atop f};z\right) \tag{2.11}$$

is given by

$$\begin{aligned} &\sum_k \frac{(a)_k(b)_k}{k!(c)_k}\frac{(d)_{n-k}(e)_{n-k}}{(n-k)!(f)_{n-k}} \\ &\quad=\sum_k \frac{(a)_k(b)_k}{k!(c)_k}\frac{(d)_n}{(1-d-n)_k}\frac{(e)_n}{(1-e-n)_k}\frac{(1-f-n)_k}{(f)_n}\frac{(-n)_k}{n!} \\ &\quad=\frac{(d)_n(e)_n}{n!(f)_n}\;{}_4F_3\left({a,b,1-f-n,-n \atop c,1-d-n,1-e-n};1\right). \end{aligned} \tag{2.12}$$

Apply Euler's transformation (2.3) on both the ${}_2F_1$'s in (2.11); one gets

$$ (1-z)^{c-a-b+f-d-e}{}_2F_1\left({c-a,c-b \atop c};z\right){}_2F_1\left({f-d,f-e \atop f};z\right) \tag{2.13}$$

When $c-a-b+f-d-e=0$ (this will correspond to the balance condition), the coefficient of z^n in (2.13) is given by (as before) :

$$ \frac{(f-d)_n(f-e)_n}{n!(f)_n}\;{}_4F_3\left({c-a,c-b,1-f-n,-n \atop c,1-f+d-n,1-f+e-n};1\right). \tag{2.14}$$

The theorem follows from equating (2.12) and (2.14), and relabelling the parameters.

The following transformation formula for a terminating ${}_3F_2$ series of unit argument is due to Sheppard, but often referred to as Whipple's ${}_3F_2$ transformation.

Corollary 2.1.

$$ {}_3F_2\left({-n, b, c \atop d, e}; 1\right) = \frac{(e-c)_n}{(e)_n}\, {}_3F_2\left({-n, d-b, c \atop d, 1+c-e-n}; 1\right). \tag{2.15}$$

Proof. Replace a by $f-a$ in (2.10). The left hand side is equal to

$$ {}_4F_3\left({-n, f-a, b, c \atop d, e, f}; 1\right). $$

Taking the limit $f \to \infty$ yields

$$ {}_3F_2\left({-n, b, c \atop d, e}; 1\right). $$

The right hand side of (2.10) reads :

$$ \frac{(e-c)_n (f-c)_n}{(e)_n (f)_n}\, {}_4F_3\left({-n, d-f+a, d-b, c \atop d, 1+c-f-n, 1+c-e-n}; 1\right). $$

Taking the same limit gives :

$$ \frac{(e-c)_n}{(e)_n}\, {}_3F_2\left({-n, d-b, c \atop d, 1+c-e-n}; 1\right). $$

2.3 Notes and Exercises

We have given only a selected number of summation and transformation formulas for (terminating) hypergeometric series in this section. For a systematic study, see [4], [43], or the more recent book [1].

In the present context, we shall need only the formulas deduced in this section, or transformation formulas derived from (2.10) or (2.15) by consecutive application.

Exercises

1. Show that

$$ {}_3F_2\left({a, b, c \atop d, e}; 1\right) = \frac{\Gamma(e)\Gamma(d+e-a-b-c)}{\Gamma(e-a)\Gamma(d+e-b-c)}\, {}_3F_2\left({a, d-b, d-c \atop d, d+e-b-c}; 1\right), \tag{2.16}$$

and

$$ {}_3F_2\left({a, b, c \atop d, e}; 1\right) = \frac{\Gamma(d)\Gamma(e)\Gamma(s)}{\Gamma(a)\Gamma(s+b)\Gamma(s+c)}\, {}_3F_2\left({d-a, e-a, s \atop s+b, s+c}; 1\right), \tag{2.17}$$

where $s = d+e-a-b-c$, and all series converge. To prove (2.16), let $n \to \infty$ in (2.10), keeping $f+n$ fixed. Apply (2.16) twice to find (2.17).

2. Consider (2.10). Keep n fixed, and apply (2.10) to itself a number of times. In this way one gets a group of transformations, realized by linear transformations on the parameters of the ${}_4F_3$. Show that, if one also includes the trivial transformations of permutations of numerator or denominator parameters, this "invariance group" behind the transformations is the symmetric group S_6; see [48].
3. In a similar way, show that the invariance group generated by the transformation (2.17) and trivial parameter permutations is the symmetric group S_5 [48].

3 Tensor product of two $\mathfrak{su}(2)$ representations

3.1 A realization of $\mathfrak{su}(2)$

Let $j \in \frac{1}{2}\mathbb{N}$, and consider the irreducible $\star$-representation D_j determined by the action (1.7). The following is a *model* or *realization* of $\mathfrak{su}(2)$ and its representation D_j :

- The representation space of D_j consists of all homogeneous polynomials in the variables x and y of degree $2j$. The basis vectors $e_m^{(j)}$ are given by

$$e_m^{(j)} = \frac{x^{j+m}y^{j-m}}{\sqrt{(j+m)!(j-m)!}}, \qquad (m = -j, -j+1, \ldots, j). \tag{3.1}$$

  The orthonormality of the basis [1]

$$\langle e_m^{(j)}, e_{m'}^{(j')}\rangle = \delta_{j,j'}\delta_{m,m'}$$

  is now equivalent with

$$\langle x^u y^v, x^{u'} y^{v'}\rangle = u!v!\delta_{u,u'}\delta_{v,v'}.$$

  or also, for arbitrary polynomials $P(x,y)$ and $Q(x,y)$:

$$\langle P(x,y), Q(x,y)\rangle = P(\partial_x, \partial_y)\cdot Q(x,y)|_{x=y=0},$$

  i.e. taking the constant part in $P(\partial_x,\partial_y)\cdot Q(x,y)$.
- The basis elements of $\mathfrak{su}(2)$ are represented by :

$$J_0 = \frac{1}{2}(x\partial_x - y\partial_y), \qquad J_+ = x\partial_y, \qquad J_- = y\partial_x. \tag{3.2}$$

So, in a realization the abstract elements of the representation are represented by some concrete vectors of a (function) space, and the abstract elements of the Lie algebra are represented by concrete operators acting in this (function) space, in such a way that the concrete action coincides with the abstract action (1.7).

[1] This can trivially be extended to distinct values j and j'.

3.2 Tensor product decomposition

Let $j_1, j_2 \in \frac{1}{2}\mathbb{N}$, and consider the tensor product $D_{j_1} \otimes D_{j_2}$, sometimes denoted by $(j_1) \otimes (j_2)$. A set of basis vectors of $D_{j_1} \otimes D_{j_2}$ is given by

$$e_{m_1}^{(j_1)} \otimes e_{m_2}^{(j_2)}, \qquad m_1 = -j_1, -j_1+1, \ldots, j_1; \;\; m_2 = -j_2, -j_2+1, \ldots, j_2. \tag{3.3}$$

This basis is often referred to as the "uncoupled basis". The total dimension of $D_{j_1} \otimes D_{j_2}$ is $(2j_1+1)(2j_2+1)$. This tensor product space is naturally equiped with an inner product by

$$\langle v \otimes w, v' \otimes w' \rangle = \langle v, v' \rangle \langle w, w' \rangle.$$

The action of the $\mathfrak{su}(2)$ basis elements on these vectors is determined by (1.4). In particular,

$$J_0(e_{m_1}^{(j_1)} \otimes e_{m_2}^{(j_2)}) = (m_1+m_2) e_{m_1}^{(j_1)} \otimes e_{m_2}^{(j_2)}. \tag{3.4}$$

So the maximal eigenvalue of J_0 is given by $j_1 + j_2$, and the corresponding eigenvector is $e_{j_1}^{(j_1)} \otimes e_{j_2}^{(j_2)}$. If we act consecutively by J_- on this vector, we get a string of vectors ending with the vector $e_{-j_1}^{(j_1)} \otimes e_{-j_2}^{(j_2)}$, with minimal J_0 eigenvalue $-j_1 - j_2$. In this way, one constructs a submodule of $D_{j_1} \otimes D_{j_2}$, of dimension $2(j_1+j_2)+1$. Clearly, this submodule does not coincide with the whole module $D_{j_1} \otimes D_{j_2}$, unless $j_1 = 0$ or $j_2 = 0$. In general, the module $D_{j_1} \otimes D_{j_2}$ is not irreducible, and it is our purpose to show that it is completely reducible and to determine its decomposition in irreducible components.

Let us first observe that simply by counting all possible J_0 eigenvalues in (3.4), this counting indicates that

$$D_{j_1} \otimes D_{j_2} = \sum_{j=|j_1-j_2|}^{j_1+j_2} D_j. \tag{3.5}$$

Of course, this is not yet a proof, since we need an argument to convince us that the irreducible components are again such $\star$-representations of $\mathfrak{su}(2)$.

In order to determine the complete decomposition, we shall explicitly construct the standard basis vectors of the representations D_j in (3.5). First of all, let us consider the vectors in $D_{j_1} \otimes D_{j_2}$ that are annihilated by J_+. For this construction, it is useful to work with the realization given earlier. Concretely, for the tensor product this realization is given by :

$$\begin{aligned} J_0 &= \frac{1}{2}(x_1\partial_{x_1} - y_1\partial_{y_1}) + \frac{1}{2}(x_2\partial_{x_2} - y_2\partial_{y_2}), \\ J_+ &= x_1\partial_{y_1} + x_2\partial_{y_2}, \qquad J_- = y_1\partial_{x_1} + y_2\partial_{x_2}, \end{aligned} \tag{3.6}$$

and

$$e_{m_1}^{(j_1)} \otimes e_{m_2}^{(j_2)} = \frac{x_1^{j_1+m_1} y_1^{j_1-m_1}}{\sqrt{(j_1+m_1)!(j_1-m_1)!}} \frac{x_2^{j_2+m_2} y_2^{j_2-m_2}}{\sqrt{(j_2+m_2)!(j_2-m_2)!}}. \tag{3.7}$$

The inner product is given by :

$$\langle x_1^{u_1} y_1^{v_1} x_2^{u_2} y_2^{v_2}, x_1^{u_1'} y_1^{v_1'} x_2^{u_2'} y_2^{v_2'} \rangle = u_1! v_1! u_2! v_2! \delta_{u_1,u_1'} \delta_{v_1,v_1'} \delta_{u_2,u_2'} \delta_{v_2,v_2'}. \tag{3.8}$$

In this realization it is clear that the following vectors are annihilated by J_+ :

$$x_1^{2j_1-k} x_2^{2j_2-k} (x_1 y_2 - x_2 y_1)^k; \tag{3.9}$$

clearly, these vectors belong to $D_{j_1} \otimes D_{j_2}$ for $k = 0, 1, \ldots, \min(2j_1, 2j_2)$. Furthermore, the J_0 eigenvalue of these vectors is given by $j_1 + j_2 - k$. Let us denote the corresponding normed vectors in $D_{j_1} \otimes D_{j_2}$ by $e_j^{(j_1 j_2) j}$, with $j = j_1 + j_2 - k$. Then $j = j_1 + j_2, j_1 + j_2 - 1, \ldots, |j_1 - j_2|$, and we write :

$$e_j^{(j_1 j_2) j} = c \sum_{l=0}^{k} (-1)^l \binom{k}{l} x_1^{2j_1 - l} y_1^l x_2^{2j_2 - k + l} y_2^{k-l},$$

where c is some constant to be determined. Expressing that these vectors have norm 1, using the inner product (3.8), one finds :

$$1 = c^2 \sum_{l=0}^{k} \binom{k}{l}^2 (2j_1 - l)! l! (2j_2 - k + l)! (k - l)!.$$

It is from this summation that c can be determined. Indeed, using trivial identities of the form

$$(a - l)! = \frac{a!}{(-1)^l (-a)_l}, \qquad (a + l)! = a! (a + 1)_l,$$

the sum is equal to (using Vandermonde's summation)

$$\begin{aligned}
k!^2 \sum_{l=0}^{k} \frac{(2j_1 - l)!(2j_2 - k + l)!}{l!(k - l)!} &= k!^2 \sum_{l=0}^{k} \frac{(2j_1)!(-k)_l (2j_2 - k)!(2j_2 - k + 1)_l}{l!(-2j_1)_l k!} \\
&= k!(2j_1)!(2j_2 - k)!\ {}_2F_1 \left({-k, 2j_2 - k + 1 \atop -2j_1} ; 1 \right) \\
&= k!(2j_1)!(2j_2 - k)! \frac{(-2j_1 - 2j_2 + k - 1)_k}{(-2j_1)_k} \\
&= \frac{k!(2j_1 - k)!(2j_2 - k)!(2j_1 + 2j_2 - k + 1)!}{(2j_1 + 2j_2 - 2k + 1)!}.
\end{aligned}$$

This determines c up to a sign. The convention is to choose c positive, so we have :

$$c = \sqrt{\frac{(2j_1 + 2j_2 - 2k + 1)!}{k!(2j_1 - k)!(2j_2 - k)!(2j_1 + 2j_2 - k + 1)!}}, \tag{3.10}$$

where $k = j_1 + j_2 - j$.

From the general action of $\mathfrak{su}(2)$, (1.8), one can construct all vectors of a representation D_j by acting on the highest weight vector e^j_j by powers of J_-. Thus we define

$$e_m^{(j_1j_2)j} = \sqrt{\frac{(j+m)!}{(2j)!(j-m)!}}(J_-)^{j-m}e_j^{(j_1j_2)j}, \tag{3.11}$$

or explicitly,

$$e_m^{(j_1j_2)j} = c'(y_1\partial_{x_1} + y_2\partial_{x_2})^{j-m}x_1^{2j_1-k}x_2^{2j_2-k}(x_1y_2 - x_2y_1)^k, \tag{3.12}$$

where $k = j_1 + j_2 - j$ and

$$\begin{aligned} c' &= \sqrt{\frac{(j+m)!}{(2j)!(j-m)!}}c = \sqrt{\frac{(j+m)!}{(2j)!(j-m)!}} \\ &\times \sqrt{\frac{(2j+1)!}{(j_1+j_2-j)!(j_1-j_2+j)!(-j_1+j_2+j)!(j_1+j_2+j+1)!}}. \end{aligned}$$

Let us work out (3.12) more explicitly :

$$\begin{aligned} e_m^{(j_1j_2)j} &= c' \sum_{i=0}^{j-m} \binom{j-m}{i} (y_1\partial_{x_1})^i (y_2\partial_{x_2})^{j-m-i} \\ &\times \sum_{l=0}^{j_1+j_2-j} (-1)^l \binom{j_1+j_2-j}{l} x_1^{2j_1-l} y_1^l x_2^{j_2-j_1+j+l} y_2^{j_1+j_2-j-l} \\ &= c' \sum_{i,l} (-1)^l \binom{j-m}{i} \binom{j_1+j_2-j}{l} \frac{(2j_1-l)!}{(2j_1-l-i)!} \frac{(j_2-j_1+j+l)!}{(j_2-j_1+l+m+i)!} \\ &\times x_1^{2j_1-l-i} y_1^{l+i} x_2^{j_2-j_1+l+m+i} y_2^{j_1+j_2-l-m-i}, \end{aligned} \tag{3.13}$$

or, using $r = l + i$:

$$\begin{aligned} &e_m^{(j_1j_2)j} \\ &= c'(j-m)! \sum_r \left(\sum_l (-1)^l \binom{2j_1-l}{r-l} \binom{j_2-j_1+j+l}{j+l-m-r} \binom{j_1+j_2-j}{l} \right) \\ &\quad \times x_1^{2j_1-r} y_1^r x_2^{j_2-j_1+m+r} y_2^{j_1+j_2-m-r} \\ &= c'(j-m)! \sum_r \left(\sum_l (-1)^l \binom{2j_1-l}{r-l} \binom{j_2-j_1+j+l}{j+l-m-r} \binom{j_1+j_2-j}{l} \right) \\ &\quad \times \sqrt{(2j_1-r)!r!(j_2-j_1+m+r)!(j_1+j_2-m-r)!}\, e_{j_1-r}^{(j_1)} \otimes e_{m+r-j_1}^{(j_2)}. \end{aligned}$$

Writing $m_1 = j_1 - r$, this can be cast in the following form :

$$e_m^{(j_1 j_2)j} = \sum_{m_1} C^{j_1,j_2,j}_{m_1,m-m_1,m} e^{(j_1)}_{m_1} \otimes e^{(j_2)}_{m-m_1},$$

where

$$\begin{aligned} & C^{j_1,j_2,j}_{m_1,m-m_1,m} \\ &= \sqrt{(j+m)!(j-m)!(j_1+m_1)!(j_1-m_1)!(j_2-m_1+m)!(j_2+m_1-m)!} \\ &\times \sqrt{\frac{(2j+1)}{(-j_1+j_2+j)!(j_1-j_2+j)!(j_1+j_2-j)!(j_1+j_2+j+1)!}} \\ &\times \sum_l (-1)^l \binom{2j_1-l}{j_1-m_1-l} \binom{j_2-j_1+j+l}{j-j_1-m+m_1+l} \binom{j_1+j_2-j}{l}. \end{aligned} \tag{3.14}$$

These coefficients are known as the Clebsch-Gordan coefficients of $\mathfrak{su}(2)$.

Theorem 3.1. *The tensor product $D_{j_1} \otimes D_{j_2}$ decomposes into irreducible $\star$-representations D_j of $\mathfrak{su}(2)$,*

$$D_{j_1} \otimes D_{j_2} = \bigoplus_{j=|j_1-j_2|}^{j_1+j_2} D_j.$$

An orthonormal basis of $D_{j_1} \otimes D_{j_2}$ is given by the vectors

$$e_m^{(j_1 j_2)j} = \sum_{m_1} C^{j_1,j_2,j}_{m_1,m-m_1,m} e^{(j_1)}_{m_1} \otimes e^{(j_2)}_{m-m_1},$$

where the coefficients $C^{j_1,j_2,j}_{m_1,m-m_1,m}$ are given by (3.14). The action of J_0, $J_\pm$ on the basis vectors $e_m^{(j_1 j_2)j}$ is the standard action (1.7) of the representation D_j.

Proof. By construction, the vectors $e_j^{(j_1 j_2)j}$ have norm 1; then the application of (3.11) implies that also $e_m^{(j_1 j_2)j}$ has norm 1. Using $\langle J_0 e_m^{(j_1 j_2)j}, e_{m'}^{(j_1 j_2)j'} \rangle = \langle e_m^{(j_1 j_2)j}, J_0 e_{m'}^{(j_1 j_2)j'} \rangle$ shows that vectors with $m \neq m'$ are orthogonal. If $m = m'$, and $j \neq j'$, then (assume $j' > j$) :

$$\begin{aligned} \langle e_m^{(j_1 j_2)j}, e_m^{(j_1 j_2)j'} \rangle &\sim \langle e_j^{(j_1 j_2)j}, (J_-)^{j'-j} e_{j'}^{(j_1 j_2)j'} \rangle \\ &\sim \langle (J_+)^{j'-j} e_j^{(j_1 j_2)j}, e_{j'}^{(j_1 j_2)j'} \rangle = 0. \end{aligned}$$

So all vectors $e_m^{(j_1 j_2)j}$ are orthonormal, hence independent. The number of such vectors is

$$\sum_{j=|j_1-j_2|}^{j_1+j_2} (2j+1) = (j_1+j_2-|j_1-j_2|+1)(j_1+j_2+|j_1-j_2|+1) = (2j_1+1)(2j_2+1),$$

hence they form an orthonormal basis. The rest of the theorem follows by construction.

3.3 Some expressions for Clebsch-Gordan coefficients of $\mathfrak{su}(2)$

The Clebsch-Gordan coefficients $C^{j_1,j_2,j}_{m_1,m-m_1,m}$ of $\mathfrak{su}(2)$ are given by (3.14). One can consider $C^{j_1,j_2,j}_{m_1,m_2,m}$ as a real function of six arguments from $\frac{1}{2}\mathbb{N}$; these arguments satisfy the following conditions :

(c1) (j_1, j_2, j) forms a *triad*, i.e. $-j_1 + j_2 + j$, $j_1 - j_2 + j$ and $j_1 + j_2 - j$ are nonnegative integers;
(c2) m_1 is a *projection* of j_1, i.e. $m_1 \in \{-j_1, -j_1 + 1, \ldots, j_1\}$ (and similarly : m_2 is a projection of j_2 and m is a projection of j);
(c3) $m = m_1 + m_2$.

Usually, one extends the definition by saying that $C^{j_1,j_2,j}_{m_1,m_2,m} = 0$ if one of the conditions (c1)-(c3) is not satisfied. Then one can write

$$e^{(j_1 j_2)j}_m = \sum_{m_1,m_2} C^{j_1,j_2,j}_{m_1,m_2,m} e^{(j_1)}_{m_1} \otimes e^{(j_2)}_{m_2},$$

and (the expressions given below are always under the conditions (c1)-(c3))

$$C^{j_1,j_2,j}_{m_1,m_2,m} = C' \sum_l (-1)^l \binom{2j_1 - l}{j_1 - m_1 - l} \binom{j_2 - j_1 + j + l}{j - j_1 - m_2 + l} \binom{j_1 + j_2 - j}{l},$$
$$C' = \sqrt{\frac{(2j+1)(j+m)!(j-m)!(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!}{(-j_1+j_2+j)!(j_1-j_2+j)!(j_1+j_2-j)!(j_1+j_2+j+1)!}}. \tag{3.15}$$

In (3.15), the summation is over all integer l-values for which the binomials are nonzero, i.e. for $\max(0, j_1 + m_2 - j) \leq l \leq \min(j_1 - m_1, j_1 + j_2 - j)$. If $j - j_1 - m_2 \geq 0$, (3.15) can directly be rewritten in terms of a ${}_3F_2$ series (otherwise, one should put $l' = j - j_1 - m_2 + l$ and then rewrite the series); one finds :

$$C^{j_1,j_2,j}_{m_1,m_2,m} = C' \frac{(2j_1)!(-j_1 + j_2 + j)!}{(j_1 + m_1)!(j_2 + m_2)!(j_1 - m_1)!(j - j_1 - m_2)!} \times {}_3F_2 \left({-j_1 + m_1, -j_1 - j_2 + j, -j_1 + j_2 + j + 1 \atop -2j_1, j - j_1 - m_2 + 1} ; 1 \right). \tag{3.16}$$

Different ${}_3F_2$ expressions appear when rewriting the sum in (3.15). For example, replacing l by $j_1 - m_1 - l$ in (3.15) yields,

$$C^{j_1,j_2,j}_{m_1,m_2,m} = C' \sum_l (-1)^{j_1 - m_1 - l} \binom{j_1 + m_1 + l}{l} \binom{j_2 + j - m_1 - l}{j - m - l} \binom{j_1 + j_2 - j}{j_1 - m_1 - l}. \tag{3.17}$$

Thus there are many ways to rewrite the expression for the $\mathfrak{su}(2)$ Clebsch-Gordan coefficients. Furthermore, when rewritten in terms of a ${}_3F_2$, one can

use transformation formulas to find yet other formulas. E.g., using (2.15) on the ${}_3F_2$ in (3.16), there comes, provided also $j-j_2+m_1\geq 0$,

$$
\begin{aligned}
C^{j_1,j_2,j}_{m_1,m_2,m} &= C'\frac{(-j_1+j_2+j)!(j_1-j_2+j)!}{(j_2+m_2)!(j_1-m_1)!(j-j_1-m_2)!(j-j_2+m_1)!} \\
&\times {}_3F_2\left({-j_1+m_1, -j_1-j_2+j, -j_2-m_2 \atop j-j_2+m_1+1, j-j_1-m_2+1} ; 1 \right).
\end{aligned}
\tag{3.18}
$$

Rewriting this last ${}_3F_2$ series again explicitly as a summation, one finds

$$
\begin{aligned}
&C^{j_1,j_2,j}_{m_1,m_2,m} = \Delta(j_1,j_2,j)\sqrt{(2j+1)} \\
&\times\sqrt{(j-m)!(j+m)!(j_1-m_1)!(j_1+m_1)!(j_2-m_2)!(j_2+m_2)!} \\
&\times\sum_l (-1)^l/(l!(j_1-m_1-l)!(j_1+j_2-j-l)!(j_2+m_2-l)! \\
&\quad\times(j-j_2+m_1+l)!(j-j_1-m_2+l)!),
\end{aligned}
\tag{3.19}
$$

where

$$
\Delta(j_1,j_2,j) = \sqrt{\frac{(-j_1+j_2+j)!(j_1-j_2+j)!(j_1+j_2-j)!}{(j_1+j_2+j+1)!}}. \tag{3.20}
$$

This rather symmetrical form is due to Van der Waerden and Racah. The expression is generally valid (that is, for all arguments satisfying (c1)-(c3)). The summation is over all integer l-values such that the factorials in the denominator of (3.19) are nonnegative.

Observe that for certain special values of the arguments, the single sum expression in (3.19) reduces to a single term. For example, put $m=j$, and consider the sum over l. The second factor in the denominator is $(j_1-m_1-l)!$, and the last factor is $(j-j_1-m_2+l)! = (m_1-j_1+l)!$. So the sum over l reduces to a single term (with $l=j_1-m_1$), and the corresponding Clebsch-Gordan coefficient has a *closed form expression* :

$$
\begin{aligned}
&C^{j_1,j_2,j}_{m_1,j-m_1,j} = (-1)^{j_1-m_1} \\
&\times\sqrt{\frac{(2j+1)!(j_1+j_2-j)!(j_1+m_1)!(j_2+j-m_1)!}{(j_1-m_1)!(j_2-j+m_1)!(-j_1+j_2+j)!(j_1-j_2+j)!(j_1+j_2+j+1)!}}.
\end{aligned}
\tag{3.21}
$$

Obviously, there exist many other expressions for the $\mathfrak{su}(2)$ Clebsch-Gordan coefficients. Usually these are given in terms of ${}_3F_2$ series, related through one another by one or more transformations of the type (2.15). In order to write a general summation (such as (3.19)) in terms of a ${}_3F_2$, one has to make extra assumptions on the parameters. For a complete list of such expressions (and many others), see e.g. [50]. Let us give one more form here. Applying (2.15) in yet another way on (3.16), one finds, under the assumption that $j_2-j_1+m\geq 0$,

$$
\begin{aligned}
C^{j_1,j_2,j}_{m_1,m_2,m} = (-1)^{j_1-m_1} C' & \frac{(2j_1)!(-j_1+j_2+j)!}{(j_1+m_1)!(j_1-m_1)!(j_2-j_1+m)!(j-m)!} \\
& \times {}_3F_2\left({-j_1+m_1, -j_1+j_2-j, -j_1+j_2+j+1 \atop -2j_1, -j_1+j_2+m+1} ; 1 \right).
\end{aligned} \tag{3.22}
$$

3.4 Symmetries of $\mathfrak{su}(2)$ Clebsch-Gordan coefficients

Various symmetries can be deduced for the Clebsch-Gordan coefficients. For example, making simple replacements $j_1 \leftrightarrow j_2$, $m_1 \leftrightarrow -m_2$, $m \leftrightarrow -m$ in (3.19) leads to

$$C^{j_1,j_2,j}_{m_1,m_2,m} = C^{j_2,j_1,j}_{-m_2,-m_1,-m}. \tag{3.23}$$

Replacing l by j_1+j_2-j-l in the summation of (3.19), and comparing with this expression with (3.19) again, one finds :

$$C^{j_1,j_2,j}_{m_1,m_2,m} = (-1)^{j_1+j_2-j} C^{j_1,j_2,j}_{-m_1,-m_2,-m}. \tag{3.24}$$

Combining these two leads to

$$C^{j_1,j_2,j}_{m_1,m_2,m} = (-1)^{j_1+j_2-j} C^{j_2,j_1,j}_{m_2,m_1,m}. \tag{3.25}$$

A somewhat surprising symmetry follows by the following replacements :

$$C^{j_1,j_2,j}_{m_1,m_2,m} = C^{(j_1+j_2+m_1+m_2)/2,(j_1+j_2-m_1-m_2)/2,j}_{(j_1-j_2+m_1-m_2)/2,(j_1-j_2-m_1+m_2)/2,j_1-j_2}. \tag{3.26}$$

It follows trivially from (3.19) under these replacements. One can verify that all conditions (c1)-(c3) are satisfied for the coefficient in the right hand side of (3.26).

Finally, replacing $j-j_2+m_1+l$ by l in the summation of (3.19), leads to

$$C^{j_1,j_2,j}_{m_1,m_2,m} = (-1)^{m_1-j_2+j} \sqrt{\frac{2j+1}{2j_2+1}} C^{j_1,j,j_2}_{-m_1,m,m_2}. \tag{3.27}$$

In order to express these, and other, symmetries, one often introduces the so-called $3j$-coefficient (due to Wigner)

$$\begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & -m \end{pmatrix} = \frac{(-1)^{j_1-j_2+m}}{\sqrt{2j+1}} C^{j_1,j_2,j}_{m_1,m_2,m}, \tag{3.28}$$

or equivalently, the corresponding Regge array

$$R_{3j}\begin{bmatrix} -j_1+j_2+j & j_1-j_2+j & j_1+j_2-j \\ j_1-m_1 & j_2-m_2 & j+m \\ j_1+m_1 & j_2+m_2 & j-m \end{bmatrix} = \begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & -m \end{pmatrix}. \tag{3.29}$$

In this 3×3 array, all entries are nonnegative integers such that for each row and each column the sum of the entries equals $J=j_1+j_2+j$; conversely, every 3×3 array with nonnegative integers such that all row and column sums are the same, corresponds to a Regge array or a $3j$-coefficient. The symmetries (3.23)-(3.27) are easy to describe in terms of the Regge array. They generate a complete group of 72 symmetries, described as follows :

- The Regge array is invariant for transposition.
- Under permutation of the rows (resp. columns), the Regge array remains invariant up to a sign. For cyclic permutations this sign is $+1$; for non-cyclic permutations this sign is $(-1)^J$.

See also exercise 4 of this section for an explanation of the symmetries. Each of these symmetries can be retranslated to the Clebsch-Gordan coefficients.

3.5 Orthogonality relations

By Theorem 3.1, we have two orthonormal bases for $D_{j_1} \otimes D_{j_2}$, namely the *uncoupled basis vectors*

$$e_{m_1}^{(j_1)} \otimes e_{m_2}^{(j_2)} \qquad (m_1 = -j_1, -j_1+1, \ldots, j_1; m_2 = -j_2, -j_2+1, \ldots, j_2)$$

and the *coupled basis vectors*

$$e_m^{(j_1 j_2)j} \qquad (j = j_1+j_2, j_1+j_2-1, \ldots, |j_1-j_2|; m = -j, -j+1, \ldots, j),$$

related by means of

$$e_m^{(j_1 j_2)j} = \sum_{m_1, m_2} C_{m_1,m_2,m}^{j_1,j_2,j} e_{m_1}^{(j_1)} \otimes e_{m_2}^{(j_2)}. \tag{3.30}$$

Since the matrix relating two orthonormal bases is orthogonal, one also has

$$e_{m_1}^{(j_1)} \otimes e_{m_2}^{(j_2)} = \sum_{j,m} C_{m_1,m_2,m}^{j_1,j_2,j} e_m^{(j_1 j_2)j}. \tag{3.31}$$

Corollary 3.1. *The Clebsch-Gordan coefficients satisfy the following orthogonality relations :*

$$\sum_{m_1,m_2} C_{m_1,m_2,m}^{j_1,j_2,j} C_{m_1,m_2,m'}^{j_1,j_2,j'} = \delta_{j,j'} \delta_{m,m'}. \tag{3.32}$$

$$\sum_{j,m} C_{m_1,m_2,m}^{j_1,j_2,j} C_{m_1',m_2',m}^{j_1,j_2,j} = \delta_{m_1,m_1'} \delta_{m_2,m_2'}. \tag{3.33}$$

Proof. For the first equation, use (3.30) in $\langle e_m^{(j_1 j_2)j}, e_{m'}^{(j_1 j_2)j'} \rangle = \delta_{j,j'} \delta_{m,m'}$, and then orthogonality of the uncoupled basis. In (3.32), one should think of j_1, j_2, j, j', m and m' being fixed; then the apparent double sum is easily seen to be actually only a single sum. Similarly, using (3.31) in $\langle e_{m_1}^{(j_1)} \otimes e_{m_2}^{(j_2)}, e_{m_1'}^{(j_1)} \otimes e_{m_2'}^{(j_2)} \rangle = \delta_{m_1,m_1'} \delta_{m_2,m_2'}$, one finds (3.33).

3.6 Recurrence relations for Clebsch-Gordan coefficients

Apart from orthogonality relations for the Clebsch-Gordan coefficients, one can also deduce certain recurrence relations in this context. We consider one case here. Use the short-hand notation

$$u_{j,m} = \sqrt{(j-m+1)(j+m)} = u_{j,-m+1}. \tag{3.34}$$

Acting by J_+ on the left hand side of (3.30) gives

$$J_+ e_m^{(j_1 j_2)j} = u_{j,m+1} e_{m+1}^{(j_1 j_2)j} = u_{j,m+1} \sum_{m_1,m_2} C_{m_1,m_2,m+1}^{j_1,j_2,j} e_{m_1}^{(j_1)} \otimes e_{m_2}^{(j_2)}.$$

Acting by J_+ on the right hand side of (3.30) yields

$$\begin{aligned}
&J_+ \left(\sum_{m_1,m_2} C_{m_1,m_2,m}^{j_1,j_2,j} e_{m_1}^{(j_1)} \otimes e_{m_2}^{(j_2)} \right) \\
&= \sum_{m_1,m_2} C_{m_1,m_2,m}^{j_1,j_2,j} u_{j_1,m_1+1} e_{m_1+1}^{(j_1)} \otimes e_{m_2}^{(j_2)} \\
&\qquad + \sum_{m_1,m_2} C_{m_1,m_2,m}^{j_1,j_2,j} u_{j_2,m_2+1} e_{m_1}^{(j_1)} \otimes e_{m_2+1}^{(j_2)} \\
&= \sum_{m_1,m_2} C_{m_1-1,m_2,m}^{j_1,j_2,j} u_{j_1,m_1} e_{m_1}^{(j_1)} \otimes e_{m_2}^{(j_2)} \\
&\qquad + \sum_{m_1,m_2} C_{m_1,m_2-1,m}^{j_1,j_2,j} u_{j_2,m_2} e_{m_1}^{(j_1)} \otimes e_{m_2}^{(j_2)}.
\end{aligned}$$

Hence,

$$u_{j,m+1} C_{m_1,m_2,m+1}^{j_1,j_2,j} = u_{j_1,m_1} C_{m_1-1,m_2,m}^{j_1,j_2,j} + u_{j_2,m_2} C_{m_1,m_2-1,m}^{j_1,j_2,j}. \tag{3.35}$$

This is only one example. Many other recurrence relations can be deduced.

3.7 Hahn and dual Hahn polynomials

Now we come to the relation between $\mathfrak{su}(2)$ Clebsch-Gordan coefficients and Hahn polynomials. Consider expression (3.22), and let us write

$$N = 2j_1,\ x = j_1 - m_1,\ n = j_1 - j_2 + j,\ \alpha = -j_1 + j_2 + m,\ \beta = -j_1 + j_2 - m. \tag{3.36}$$

Think of j_1, j_2 and m as being fixed numbers, with

$$j_2 - j_1 \geq m \geq 0. \tag{3.37}$$

Then m_1 can vary between $-j_1$ and j_1, and j can vary between $j_2 - j_1$ and $j_2 + j_1$. In terms of the new variables, this means : N is a fixed nonnegative

integer, $\alpha \geq 0$ and $\beta \geq 0$ are fixed; the quantities x and n are nonnegative integers with $0 \leq x \leq N$ and $0 \leq n \leq N$. The ${}_3F_2$ series appearing in (3.22) is then of the following form :

$$Q_n(x;\alpha,\beta,N) = {}_3F_2\left({-x,-n,1+n+\alpha+\beta \atop -N,1+\alpha} ;1\right). \tag{3.38}$$

For fixed $\alpha,\beta \geq 0$ and a nonnegative integer N, this expression is a polynomial of degree n in the variable x, for all $0 \leq n \leq N$. This polynomial is the *Hahn polynomial*. In general, it is defined for $N \in \mathbb{N}$, and $\alpha,\beta \notin \{-1,-2,\ldots,-N\}$.

In terms of the currently introduced variables, (3.22) becomes

$$C^{j_1,j_2,j}_{m_1,m_2,m} = \\ (-1)^x \left(\frac{(\alpha+\beta+2n+1)(\alpha+n)!(\alpha+x)!(\beta+N-x)!(\alpha+\beta+n)!}{n!(N-n)!x!(N-x)!(N+\alpha+\beta+n+1)!(\beta+n)!} \right)^{1/2} \\ \times \frac{N!}{\alpha!}\, Q_n(x;\alpha,\beta,N).$$

The orthogonality relation (3.32) with $m = m'$ can now be written in the following form :

$$\sum_{x=0}^{N} \frac{(1+\alpha)_x}{x!}\frac{(1+\beta)_{N-x}}{(N-x)!} Q_n(x;\alpha,\beta,N)Q_{n'}(x;\alpha,\beta,N) = h_n(\alpha,\beta,N)\delta_{n,n'}, \tag{3.39}$$

where

$$h_n(\alpha,\beta,N) = \frac{n!(N-n)!}{N!^2}\frac{(1+\beta)_n}{(1+\alpha)_n}\frac{(1+\alpha+\beta+n)_{N+1}}{(1+\alpha+\beta+2n)}. \tag{3.40}$$

This is the (discrete) orthogonality relation for Hahn polynomials. It follows directly from the orthogonality of $\mathfrak{su}(2)$ Clebsch-Gordan coefficients, at least in the case that α and β are nonnegative integers. However, the final expressions (3.39) and (3.40) are rational in the parameters α and β. So the relation is valid for all values of α and β provided the expressions are well defined. In other words,

Theorem 3.2. *Let α, β and N be given, with N a nonnegative integer and $\alpha,\beta \notin \{-1,-2,\ldots,-N\}$. Then the Hahn polynomials $Q_n(x;\alpha,\beta,N)$, defined in (3.38), satisfy the orthogonality relation (3.39).*

We also see from (3.39) that if $\alpha,\beta > -1$ the weight function on the left hand side is positive, while if $\alpha,\beta < -N$ it becomes positive after multiplication with $(-1)^N$. So in these cases (3.39) gives a system of polynomials orthogonal with respect to a positive measure.

In a similar way, one can under the same conditions write the orthogonality relation (3.33) as

$$\sum_{n=0}^{N} \frac{N!^2}{n!(N-n)!} \frac{(1+\alpha)_n}{(1+\beta)_n} \frac{(1+\alpha+\beta+2n)}{(1+\alpha+\beta+n)_{N+1}} Q_n(x;\alpha,\beta,N) Q_n(x';\alpha,\beta,N)$$
$$= \frac{x!(N-x)!}{(1+\alpha)_x(1+\beta)_{N-x}} \delta_{x,x'}. \tag{3.41}$$

Changing the role of x and n, this gives the orthogonality for the *dual Hahn polynomials*. Let

$$R_n(\lambda(x);\alpha,\beta,N) \equiv Q_x(n;\alpha,\beta,N) = {}_3F_2\left(\begin{matrix} -n, -x, 1+x+\alpha+\beta \\ -N, 1+\alpha \end{matrix} ; 1 \right). \tag{3.42}$$

As before, $N \in \mathbb{N}$, $0 \le n, x \le N$ and $\alpha, \beta \notin \{-1, -2, \ldots, -N\}$. In this case, it is easy to verify that $R_n(\lambda(x);\alpha,\beta,N)$ is a polynomial of degree n in the variable

$$\lambda(x) = x(x+\alpha+\beta+1).$$

Now (3.41) yields

Theorem 3.3. *The dual Hahn polynomials satisfy the orthogonality relation*

$$\sum_{x=0}^{N} w(x;\alpha,\beta,N) R_n(\lambda(x);\alpha,\beta,N) R_{n'}(\lambda(x);\alpha,\beta,N) = h'_n(\alpha,\beta,N)\delta_{n,n'},$$

where

$$w(x;\alpha,\beta,N) = \frac{(-1)^x N!(-N)_x(1+\alpha)_x(1+\alpha+\beta+2x)}{x!(1+\beta)_x(1+\alpha+\beta+x)_{N+1}},$$

and

$$h'_n(\alpha,\beta,N) = \frac{n!(N-n)!}{(1+\alpha)_n(1+\beta)_{N-n}}.$$

3.8 Recurrence relations

The various recurrence relations that can be deduced for Clebsch-Gordan coefficients in the context of representation theory, lead to relations for the Hahn or dual Hahn polynomials. Let us give one example. In (3.35) we have deduced from the J_+ action that

$$u_{j,m} C^{j_1,j_2,j}_{m_1,m_2,m} = u_{j_1,m_1} C^{j_1,j_2,j}_{m_1-1,m_2,m-1} + u_{j_2,m_2} C^{j_1,j_2,j}_{m_1,m_2-1,m-1}, \tag{3.43}$$

where $u_{j,m}$ is given in (3.34). Similarly, from the action of J_- one finds :

$$u_{j,-m} C^{j_1,j_2,j}_{m_1,m_2,m} = u_{j_1,-m_1} C^{j_1,j_2,j}_{m_1+1,m_2,m+1} + u_{j_2,-m_2} C^{j_1,j_2,j}_{m_1,m_2+1,m+1}. \tag{3.44}$$

Combining the two, one gets :

$$
\begin{aligned}
&u_{j,m}u_{j,-m+1}C^{j_1,j_2,j}_{m_1,m_2,m} = \\
&\quad u_{j_1,m_1}[u_{j_1,-m_1+1}C^{j_1,j_2,j}_{m_1,m_2,m} + u_{j_2,-m_2}C^{j_1,j_2,j}_{m_1-1,m_2+1,m}] \\
&\qquad +u_{j_2,m_2}[u_{j_1,-m_1}C^{j_1,j_2,j}_{m_1+1,m_2-1,m} + u_{j_2,-m_2+1}C^{j_1,j_2,j}_{m_1,m_2,m}].
\end{aligned}
$$

This can be rewritten as

$$
\begin{aligned}
&u^2_{j,m}C^{j_1,j_2,j}_{m_1,m_2,m} = u_{j_1,m_1}u_{j_2,-m_2}C^{j_1,j_2,j}_{m_1-1,m_2+1,m} \\
&+(u^2_{j_1,m_1} + u^2_{j_2,m_2})C^{j_1,j_2,j}_{m_1,m_2,m} + u_{j_2,m_2}u_{j_1,-m_1}C^{j_1,j_2,j}_{m_1+1,m_2-1,m}.
\end{aligned}
$$

Using the relation with dual Hahn polynomials :

$$
\begin{aligned}
&C^{j_1,j_2,j}_{m_1,m_2,m} = \\
&(-1)^n \left(\frac{(\alpha+\beta+2x+1)(\alpha+x)!(\alpha+n)!(\beta+N-n)!(\alpha+\beta+x)!}{x!(N-x)!n!(N-n)!(N+\alpha+\beta+x+1)!(\beta+x)!} \right)^{1/2} \\
&\qquad \times \frac{N!}{\alpha!}\, R_n(\lambda(x);\alpha,\beta,N),
\end{aligned}
$$

where

$$
N=2j_1,\ n=j_1-m_1,\ x=j_1-j_2+j,\ \alpha=-j_1+j_2+m,\ \beta=-j_1+j_2-m,
$$

this yields the classical recurrence relation for dual Hahn polynomials :

$$
\lambda(x)R_n(\lambda(x)) = A_nR_{n+1}(\lambda(x)) - (A_n+C_n)R_n(\lambda(x)) + C_nR_{n-1}(\lambda(x)),
$$

where $R_n(\lambda(x)) \equiv R_n(\lambda(x);\alpha,\beta,N)$ and

$$
A_n = (n-N)(1+\alpha+n), \qquad C_n = n(n-\beta-N-1).
$$

3.9 Tensor product of $\mathfrak{su}(1,1)$ representations

Finally, recall the irreducible $\star$-representations of $\mathfrak{su}(1,1)$, determined in (1.11). For convenience, let us rename the label λ by k. So k is a positive real number, and the representation (1.11) is usually referred to as *the positive discrete series representation* with label k of $\mathfrak{su}(1,1)$ (this is because $\mathfrak{su}(1,1)$ also has irreducible $\star$-representations of a different type, e.g. negative discrete series representations, continuous series, ...), and we denote it by $\mathcal{D}_k$. An orthonormal basis of $\mathcal{D}_k$ is given by the vectors $e_n^{(k)}$ $(n=0,1,\ldots)$, and the $\mathfrak{su}(1,1)$ action is :

$$
\begin{aligned}
J_0 e_n^{(k)} &= (k+n)e_n^{(k)}, \\
J_+ e_n^{(k)} &= \sqrt{(n+1)(2k+n)}\; e_{n+1}^{(k)}, \\
J_- e_n^{(k)} &= -\sqrt{n(2k+n-1)}\; e_{n-1}^{(k)}.
\end{aligned}
\tag{3.45}
$$

The technique followed here to determine tensor product decompositions and Clebsch-Gordan coefficients for $\mathfrak{su}(2)$ can be copied completely to the case of positive discrete series representations of $\mathfrak{su}(1,1)$. The following is a useful realization for $\mathfrak{su}(1,1)$ and for the representation $\mathcal{D}_k$ (with $k > 1/2$) :

$$e_n^{(k)} = \sqrt{\frac{(2k)_n}{n!}} z^n; \tag{3.46}$$

the representation space is the Hilbert space of analytic functions $f(z)$ ($z \in \mathbb{C}$) on the unit disc $|z| < 1$ with inner product

$$\langle f_1, f_2 \rangle = \frac{2k-1}{\pi} \iint_{|z|<1} f_1(z)\overline{f_2(z)}(1-|z|^2)^{2k-2} dxdy, \qquad (z = x+iy). \tag{3.47}$$

One can verify that $\langle e_n^{(k)}, e_m^{(k)} \rangle = \delta_{m,n}$. The realization of the $\mathfrak{su}(1,1)$ basis elements reads as follows :

$$J_0 = z\frac{d}{dz} + k, \qquad J_- = -\frac{d}{dz}, \qquad J_+ = z^2\frac{d}{dz} + 2kz. \tag{3.48}$$

It is easy to verify that the action of these operators on the basis (3.46) is indeed the same as in (3.45), and that all $\star$-conditions are satisfied.

To determine the tensor product decomposition $\mathcal{D}_{k_1} \otimes \mathcal{D}_{k_2}$, one can use this realization. By the tensor product rule, the basis functions $e_{n_1}^{(k_1)} \otimes e_{n_2}^{(k_2)}$ are, up to a constant, given by $z_1^{n_1} z_2^{n_2}$; and the operators are

$$\begin{aligned} &J_0 = z_1\partial_{z_1} + z_2\partial_{z_2} + k_1 + k_2, \qquad J_- = -\partial_{z_1} - \partial_{z_2}, \\ &J_+ = z_1^2\partial_{z_1} + z_2^2\partial_{z_2} + 2k_1z_1 + 2k_2z_2. \end{aligned} \tag{3.49}$$

Clearly, the vectors annihilated by J_- are again easy to construct; up to a constant these are just $(z_1 - z_2)^j$, with $j \in \mathbb{N}$. These vectors are the analogues of (3.9). Normalizing them yields vectors $e_0^{(k_1,k_2)k_1+k_2+j}$ which are annihilated by J_- and have J_0 eigenvalue $k_1 + k_2 + j \equiv k$. Then one can act by positive powers of J_+ on these vectors using the general expression $J_+^n e_0^{(k)} = \sqrt{n!(2k)_n} e_n^{(k)}$ [this is the analogue of (3.11)]. The final result is a complete decomposition of the tensor product, and an explicit expression for the Clebsch-Gordan coefficients. One finds :

$$\mathcal{D}_{k_1} \otimes \mathcal{D}_{k_2} = \sum_{j=0}^{\infty} \mathcal{D}_{k_1+k_2+j}, \tag{3.50}$$

and the $\mathfrak{su}(1,1)$ Clebsch-Gordan coefficients for these positive discrete series, i.e. for $k = k_1 + k_2 + j$, ($j \in \mathbb{N}$), $k + n = k_1 + n_1 + k_2 + n_2$,

$$e_n^{(k_1,k_2)k} = \sum_{n_1,n_2} C_{n_1,n_2,n}^{k_1,k_2,k} e_{n_1}^{(k_1)} \otimes e_{n_2}^{(k_2)}, \tag{3.51}$$

are given by (with $j=k-k_1-k_2$ and $n=n_1+n_2-j$) :

$$C_{n_1,n_2,n}^{k_1,k_2,k} = \left[\frac{(2k_1)_{n_1}(2k_2)_{n_2}(2k_1)_j}{n!n_1!n_2!j!(2k_1+2k_2+2j)_n(2k_2)_j(2k_1+2k_2+j-1)_j} \right]^{1/2} \times (j+n)!\ {}_3F_2\left({2k_1+2k_2+j-1, -n_1, -j \atop 2k_1, -n-j} ; 1 \right). \tag{3.52}$$

Orthogonality relations for $\mathfrak{su}(1,1)$ Clebsch-Gordan coefficients are again equivalent with the orthogonality of Hahn or dual Hahn polynomials.

3.10 Notes and Exercises

Clebsch-Gordan coefficients (sometimes called Wigner coefficients) were introduced in [31]. Wigner [53] derived a ${}_3F_2$ type expression for these coefficients, and introduced the symmetric form for the $3j$-coefficients [52]. The theory of Clebsch-Gordan coefficients was highly influenced by the quantum theory of angular momentum. In this theory the $3j$-coefficients play a crucial role; see also the papers in [8] and the survey in [44].

The realization in section 3.1 is often referred to as Schwinger's realization [41], since it is closely related to a realization in terms of creation and annihilation operators. The tensor product decomposition in section 3.2 follows the technique used in many physics books; in their terminology, taking tensor products of $\mathfrak{su}(2)$ representations corresponds to coupling of angular momenta. If (j_1,m_1) denotes the angular momentum labels of the first particle (j_1 its total value, and m_1 its projection along a chosen axis), and (j_2,m_2) those of the second particle, then (j,m) gives the angular momentum labels of the system consisting of these two particles.

The various expressions for $\mathfrak{su}(2)$ Clebsch-Gordan coefficients were derived by Wigner [53], Racah [33] and Majumdar [30]. The relations between these expressions, and their domain of definition, were studied in detail by Srinivasa Rao [35], [36]; see also [7, volume 9, special topic 11].

The classical symmetries of $\mathfrak{su}(2)$ Clebsch-Gordan coefficients (corresponding to equation (3.23), (3.24), (3.25) and (3.27)) were already known to Wigner [52]; he used them to define the more symmetric $3j$-coefficient. It was only until much later that Regge discovered the extra symmetries, as in (3.26), now known as the Regge symmetries [37].

Hahn polynomials appeared already in the work of Chebyshev; see also [24]. Their relation with $\mathfrak{su}(2)$ Clebsch-Gordan coefficients was apparently known to some people but appeared in the literature only in 1981 [27]; see also [45], [51]. For a nice introduction to this relationship in the context of representations of the Lie group $SU(2)$, see [27] (here we have approached it purely on the level of the Lie algebra representations). A summary of properties of Hahn polynomials and relations with other orthogonal polynomials is given in [25].

An impressive collection of formulas for $3j$-coefficients (and many other objects related to quantum theory of angular momentum) is found in [50]. To my knowledge, this is the most complete and accurate book as far as formulas are concerned.

Clebsch-Gordan coefficients for positive discrete series of $\mathfrak{su}(1,1)$ are given in [51]; their origin goes back to the work of Biedenharn and others [18], [46], [34]. For an overview of actual decompositions of tensor products of unitary $\mathfrak{su}(1,1)$ representations in general, see [39].

Exercises

1. Verify that the operators (3.2) satisfy (1.3), and that the action of these operators on (3.1) coincides with (1.7).
2. Following (3.7) and (3.30), one can write

$$e_m^{(j_1 j_2)j} = \sum_{m_1,m_2} C^{j_1,j_2,j}_{m_1,m_2,m} \frac{x_1^{j_1+m_1} y_1^{j_1-m_1} x_2^{j_2+m_2} y_2^{j_2-m_2}}{\sqrt{(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!}}.$$

   Multiply both sides of this equation by $x^{j+m}y^{j-m}/\sqrt{(j+m)!(j-m)!}$, and sum over all m-values. To work out the left hand side, use e.g. (3.13). Show that this leads to the following identity :

$$\begin{aligned} &\frac{c}{\sqrt{(2j)!}}(x_1y_2-x_2y_1)^{j_1+j_2-j}(x_1x+y_1y)^{j_1-j_2+j}(x_2x+y_2y)^{-j_1+j_2+j} \\ &= \sum_{m_1,m_2,m} C^{j_1,j_2,j}_{m_1,m_2,m} \\ &\times \frac{x_1^{j_1+m_1} y_1^{j_1-m_1} x_2^{j_2+m_2} y_2^{j_2-m_2} x^{j+m} y^{j-m}}{\sqrt{(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!(j+m)!(j-m)!}}, \end{aligned} \tag{3.53}$$

   where c is given in (3.10). This can be seen as a generating function for the $\mathfrak{su}(2)$ Clebsch-Gordan coefficients.
3. Rewrite (3.53) in terms of $3j$-coefficients to find :

$$\begin{aligned} &\frac{(x_2y_3-x_3y_2)^{-j_1+j_2+j_3}(x_3y_1-x_1y_3)^{j_1-j_2+j_3}(x_1y_2-x_2y_1)^{j_1+j_2-j_3}}{\sqrt{(-j_1+j_2+j_3)!(j_1-j_2+j_3)!(j_1+j_2-j_3)!}} \\ &= \sqrt{(J+1)!} \sum_{m_1,m_2,m_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \\ &\times \frac{x_1^{j_1+m_1} y_1^{j_1-m_1} x_2^{j_2+m_2} y_2^{j_2-m_2} x_3^{j_3+m_3} y_3^{j_3-m_3}}{\sqrt{(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!(j_3+m_3)!(j_3-m_3)!}}, \end{aligned} \tag{3.54}$$

   where $J = j_1 + j_2 + j_3$.
4. In (3.54), multiply both sides by $z_1^{-j_1+j_2+j_3} z_2^{j_1-j_2+j_3} z_3^{j_1+j_2-j_3}$, and sum over all values of j_1, j_2 and j_3 with $j_1+j_2+j_3 = J$ fixed, with (j_1, j_2, j_3) a triad. Show that this leads to the following beautiful determinant formula :

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}^J = (-1)^J J! \sqrt{(J+1)!}$$

$$\times \sum_R R \frac{x_1^{R_{11}} x_2^{R_{12}} x_3^{R_{13}} y_1^{R_{21}} y_2^{R_{22}} y_3^{R_{23}} z_1^{R_{31}} z_2^{R_{32}} z_3^{R_{33}}}{\sqrt{R_{11}!R_{12}!R_{13}!R_{21}!R_{22}!R_{23}!R_{31}!R_{32}!R_{33}!}}.$$

Herein, the sum is over all Regge arrays (3.29), i.e. over all 3×3 magic squares of non-negative integers, with $\sum_i R_{ij} = J$ for all j, and $\sum_j R_{ij} = J$ for all i. This formula is due to Regge [37]; the symmetries of the $3j$-coefficient can easily be deduced from here.

5. Verify that the $\mathfrak{su}(1,1)$ Clebsch-Gordan coefficients (3.52) satisfy :

$$C^{k_2,k_1,k}_{n_2,n_1,n} = (-1)^{k_1+k_2-k} C^{k_1,k_2,k}_{n_1,n_2,n}. \tag{3.55}$$

6. The realization in section 3.1 is on polynomials in two variables, whereas that in section 3.9 is on analytic functions in one variable. Also for $\mathfrak{su}(2)$ there is a realization on polynomials in one variable of degree at most $2j$. Show that this realization is given by :

$$e^{(j)}_m = \sqrt{\frac{(2j)!}{(j+m)!(j-m)!}}\, z^{j+m}, \qquad (m=-j,\ldots,j), \tag{3.56}$$

with

$$J_0 = z\frac{d}{dz} - j, \qquad J_- = \frac{d}{dz}, \qquad J_+ = -z^2\frac{d}{dz} + 2jz. \tag{3.57}$$

Furthermore, show that the inner product $\langle e^{(j)}_m, e^{(j)}_{m'}\rangle = \delta_{m,m'}$ follows from the integral representation $(z = x+iy)$

$$\langle f_1, f_2\rangle = \frac{2j+1}{\pi} \iint f_1(z)\overline{f_2(z)}(1+|z|^2)^{-2j-2} dx dy, \tag{3.58}$$

where the integration is on the whole complex plane (more precisely, on the Riemann sphere).

4 Tensor product of three $\mathfrak{su}(2)$ representations and Racah coefficients

4.1 Definition of the Racah coefficient

Let us consider the tensor product of three irreducible $\star$-representations of $\mathfrak{su}(2)$,

$$D_{j_1} \otimes D_{j_2} \otimes D_{j_3} = (D_{j_1} \otimes D_{j_2}) \otimes D_{j_3} = D_{j_1} \otimes (D_{j_2} \otimes D_{j_3}). \tag{4.1}$$

Clearly, a basis for this tensor product is given by

$$e^{(j_1)}_{m_1} \otimes e^{(j_2)}_{m_2} \otimes e^{(j_3)}_{m_3},$$

where m_i is a projection of j_i. These are the $(2j_1+1)(2j_2+1)(2j_3+1)$ orthonormal basis vectors of the *uncoupled basis*.

In order to decompose the actual tensor product $D_{j_1} \otimes D_{j_2} \otimes D_{j_3}$ into irreducible $\mathfrak{su}(2)$ representations, one can work as follows. First, decompose $D_{j_1} \otimes D_{j_2}$ into irreducibles, and then decompose the tensor product of each such irreducible with D_{j_3}. Thus we have

$$(\sum_{j_{12}=|j_1-j_2|}^{j_1+j_2} D_{j_{12}}) \otimes D_{j_3} = \sum_{j_{12}=|j_1-j_2|}^{j_1+j_2} \sum_{j=|j_{12}-j_3|}^{j_{12}+j_3} D_j.$$

It is clear that now multiplicities will arise in this tensor product decomposition, e.g. if $j_1 = 1$, $j_2 = 2$ and $j_3 = 2$ then

$$\begin{aligned}\sum_{j_{12}=1}^{3} \sum_{j=|j_{12}-2|}^{j_{12}+2} D_j &= (D_1 + D_2 + D_3) + (D_0 + D_1 + D_2 + D_3 + D_4) \\ &\quad +(D_1 + D_2 + D_3 + D_4 + D_5) \\ &= D_0 + 3D_1 + 3D_2 + 3D_3 + 2D_4 + D_5.\end{aligned}$$

Secondly, one can also work differently and first decompose $D_{j_2} \otimes D_{j_3}$ into irreducibles :

$$D_{j_1} \otimes (\sum_{j_{23}=|j_2-j_3|}^{j_2+j_3} D_{j_{23}}) = \sum_{j_{23}=|j_2-j_3|}^{j_2+j_3} \sum_{j=|j_1-j_{23}|}^{j_1+j_{23}} D_j.$$

The final decomposition is of course the same; for the above example this becomes :

$$\begin{aligned}\sum_{j_{23}=0}^{4} \sum_{j=|1-j_{23}|}^{1+j_{23}} D_j &= (D_1) + (D_0 + D_1 + D_2) + (D_1 + D_2 + D_3) \\ &\quad +(D_2 + D_3 + D_4) + (D_3 + D_4 + D_5) \\ &= D_0 + 3D_1 + 3D_2 + 3D_3 + 2D_4 + D_5.\end{aligned}$$

In order to introduce an orthonormal basis for the irreducible components of the tensor product, one has to use an extra label to resolve the multiplicity problem. This extra label is simply provided by the representation label of the *intermediate coupling*, i.e. by j_{12} or j_{23}. So one can immediately define two sets of orthonormal basis vectors for the irreducible components of (4.1),

$$\begin{aligned}e_m^{((j_1j_2)j_{12}j_3)j} &= \sum_{m_{12},m_3} C^{j_{12},j_3,j}_{m_{12},m_3,m} e^{(j_1j_2)j_{12}}_{m_{12}} \otimes e^{(j_3)}_{m_3} \qquad (4.2)\\ &= \sum_{\substack{m_1,m_2,m_3\\ m_1+m_2+m_3=m}} C^{j_1,j_2,j_{12}}_{m_1,m_2,m_{12}} C^{j_{12},j_3,j}_{m_{12},m_3,m}\, e^{(j_1)}_{m_1} \otimes e^{(j_2)}_{m_2} \otimes e^{(j_3)}_{m_3},\end{aligned}$$

and

$$
\begin{aligned}
e_m^{(j_1(j_2j_3)j_{23})j} &= \sum_{m_1,m_{23}} C^{j_1,j_{23},j}_{m_1,m_{23},m} e^{(j_1)}_{m_1} \otimes e^{(j_2j_3)j_{23}}_{m_{23}} \\
&= \sum_{\substack{m_1,m_2,m_3 \\ m_1+m_2+m_3=m}} C^{j_2,j_3,j_{23}}_{m_2,m_3,m_{23}} C^{j_1,j_{23},j}_{m_1,m_{23},m}\; e^{(j_1)}_{m_1} \otimes e^{(j_2)}_{m_2} \otimes e^{(j_3)}_{m_3}.
\end{aligned} \tag{4.3}
$$

The transition from one orthonormal basis to a second one yields an orthogonal matrix. Let us denote the matrix transforming (4.2) into (4.3) by U. Its matrix elements are given by

$$
\langle e_m^{(j_1(j_2j_3)j_{23})j}, e_{m'}^{((j_1j_2)j_{12}j_3)j'} \rangle.
$$

Using the Casimir operator,

$$
\langle C e_m^{(j_1(j_2j_3)j_{23})j}, e_{m'}^{((j_1j_2)j_{12}j_3)j'} \rangle = \langle e_m^{(j_1(j_2j_3)j_{23})j}, C e_{m'}^{((j_1j_2)j_{12}j_3)j'} \rangle,
$$

it follows immediately that this element is zero if $j' \neq j$. Similarly, replacing C by J_0 implies that the element is also zero when $m' \neq m$. Next, from the action of J_+, one finds

$$
\begin{aligned}
&\langle J_+ e_{m-1}^{(j_1(j_2j_3)j_{23})j}, e_m^{((j_1j_2)j_{12}j_3)j} \rangle \\
=&\sqrt{(j-m+1)(j+m)} \langle e_m^{(j_1(j_2j_3)j_{23})j}, e_m^{((j_1j_2)j_{12}j_3)j} \rangle \\
=&\langle e_{m-1}^{(j_1(j_2j_3)j_{23})j}, J_- e_m^{((j_1j_2)j_{12}j_3)j} \rangle \\
=&\sqrt{(j+m)(j-m+1)} \langle e_{m-1}^{(j_1(j_2j_3)j_{23})j}, e_{m-1}^{((j_1j_2)j_{12}j_3)j} \rangle.
\end{aligned}
$$

So the matrix element $\langle e_m^{(j_1(j_2j_3)j_{23})j}, e_m^{((j_1j_2)j_{12}j_3)j} \rangle$ is independent of m. Let us therefore denote

$$
\langle e_m^{(j_1(j_2j_3)j_{23})j}, e_{m'}^{((j_1j_2)j_{12}j_3)j'} \rangle = \delta_{j,j'} \delta_{m,m'} U^{j_1,j_2,j_{12}}_{j_3,j,j_{23}}. \tag{4.4}
$$

The coefficients $U^{j_1,j_2,j_{12}}_{j_3,j,j_{23}}$ are called the *Racah coefficients*. So, we can write

$$
e_m^{((j_1j_2)j_{12}j_3)j} = \sum_{j_{23}} U^{j_1,j_2,j_{12}}_{j_3,j,j_{23}} e_m^{(j_1(j_2j_3)j_{23})j}, \tag{4.5}
$$

and vice versa, since U is an orthogonal matrix,

$$
e_m^{(j_1(j_2j_3)j_{23})j} = \sum_{j_{12}} U^{j_1,j_2,j_{12}}_{j_3,j,j_{23}} e_m^{((j_1j_2)j_{12}j_3)j}. \tag{4.6}
$$

The orthogonality of the matrix is also expressed by :

$$
\sum_{j_{12}} U^{j_1,j_2,j_{12}}_{j_3,j,j_{23}} U^{j_1,j_2,j_{12}}_{j_3,j,j'_{23}} = \delta_{j_{23},j'_{23}}, \tag{4.7}
$$

$$
\sum_{j_{23}} U^{j_1,j_2,j_{12}}_{j_3,j,j_{23}} U^{j_1,j_2,j'_{12}}_{j_3,j,j_{23}} = \delta_{j_{12},j'_{12}}. \tag{4.8}
$$

4.2 The $6j$-coefficient

An appropriate expression for the Racah coefficient follows from (4.4), (4.2) and (4.3) :

$$U^{j_1,j_2,j_{12}}_{j_3,j,j_{23}} = \sum_{\substack{m_1,m_2,m_3 \\ m_1+m_2+m_3=m}} C^{j_1,j_2,j_{12}}_{m_1,m_2,m_{12}} C^{j_{12},j_3,j}_{m_{12},m_3,m} C^{j_2,j_3,j_{23}}_{m_2,m_3,m_{23}} C^{j_1,j_{23},j}_{m_1,m_{23},m}. \tag{4.9}$$

Herein, m is an arbitrary but fixed projection of j. The sum is over m_1, m_2 and m_3 such that $m_1+m_2+m_3=m$; m_{12} stands for m_1+m_2 and m_{23} for m_2+m_3. So this is a double sum over the product of four Clebsch-Gordan coefficients. This is clearly a rather complicated object. We shall see how to simplify this expression later.

The current expression, however, is useful to deduce some symmetry properties of the Racah coefficient. First of all, summing over the $(2j+1)$ possible values for m in (4.9) yields

$$(2j+1)U^{j_1,j_2,j_{12}}_{j_3,j,j_{23}} = \sum C^{j_1,j_2,j_{12}}_{m_1,m_2,m_{12}} C^{j_{12},j_3,j}_{m_{12},m_3,m} C^{j_2,j_3,j_{23}}_{m_2,m_3,m_{23}} C^{j_1,j_{23},j}_{m_1,m_{23},m}, \tag{4.10}$$

where the sum is now over all possible m-values (taking into account the conditions (c1)-(c3) for Clebsch-Gordan coefficients). Changing names of the variables gives

$$(2j_{23}+1)U^{j_1,j_{12},j_2}_{j_3,j_{23},j} = \sum C^{j_1,j_{12},j_2}_{m_1,m_{12},m_2} C^{j_2,j_3,j_{23}}_{m_2,m_3,m_{23}} C^{j_{12},j_3,j}_{m_{12},m_3,m} C^{j_1,j,j_{23}}_{m_1,m,m_{23}}.$$

Applying (3.27) to the first and last C-coefficient yields

$$\begin{aligned}(2j_{23}+1)U^{j_1,j_{12},j_2}_{j_3,j_{23},j} &= (-1)^{j_2-j_{12}+j-j_{23}}\sqrt{\frac{(2j_2+1)(2j_{23}+1)}{(2j_{12}+1)(2j+1)}} \\ \times \sum & C^{j_1,j_2,j_{12}}_{m_1,m_2,m_{12}} C^{j_2,j_3,j_{23}}_{m_2,m_3,m_{23}} C^{j_{12},j_3,j}_{m_{12},m_3,m} C^{j_1,j_{23},j}_{m_1,m_{23},m}.\end{aligned}$$

Thus, we find :

$$U^{j_1,j_{12},j_2}_{j_3,j_{23},j} = (-1)^{j_2-j_{12}+j-j_{23}}\sqrt{\frac{(2j_2+1)(2j+1)}{(2j_{12}+1)(2j_{23}+1)}}U^{j_1,j_2,j_{12}}_{j_3,j,j_{23}}. \tag{4.11}$$

This is only one example of a symmetry. Many others can be deduced from (4.10) and the symmetry properties of Clebsch-Gordan coefficients.

In this context, it is customary to introduce the so-called $6j$*-coefficient*,

$$\begin{Bmatrix} a & b & c \\ d & e & f \end{Bmatrix} = (-1)^{a+b+d+e}\frac{U^{a,b,c}_{d,e,f}}{\sqrt{(2c+1)(2f+1)}}, \tag{4.12}$$

where (a,b,c), (d,e,c), (d,b,f) and (a,e,f) are triads. Then the $6j$-coefficient is invariant under any permutation of its columns, or under the interchange of the upper and lower arguments in each of any two columns.

Even more, one can also use the Regge symmetries of the Clebsch-Gordan coefficients, and obtain similar symmetries for the $6j$-coefficient. In order to describe these, let the Regge array for the $6j$-coefficient be defined as the 3×4 array

$$\begin{Bmatrix} a & b & c \\ d & e & f \end{Bmatrix} = R_{6j} \begin{bmatrix} d+e-c & b+d-f & a+e-f & a+b-c \\ d+f-b & c+d-e & a+c-b & a+f-e \\ e+f-a & b+c-a & c+e-d & b+f-d \end{bmatrix}.$$

Then the value of the Regge array is invariant under any permutation of its rows or columns. Note that the arguments of the Regge array are such that all entries are nonnegative integers, and the differences between corresponding elements of rows (resp. columns) are constant. Conversely, every 3×4 array of nonnegative integers with this property corresponds to a Regge array, or a $6j$-coefficient.

4.3 Expressions for the Racah coefficient

Racah was the first to simplify the expression for $U_{j_3,j,j_{23}}^{j_1,j_2,j_{12}}$, leading finally to a single sum expression. We shall follow here the method of Vilenkin and Klimyk [51] to obtain Racah's expression.

From (4.6), one finds, using (4.2) and (4.3),

$$\begin{aligned} & C_{m_2,m_3,m_2+m_3}^{j_2,j_3,j_{23}} C_{m_1,m_2+m_3,m_1+m_2+m_3}^{j_1,j_{23},j} \\ & \qquad = \sum_{j_{12}} U_{j_3,j,j_{23}}^{j_1,j_2,j_{12}} C_{m_1,m_2,m_{12}}^{j_1,j_2,j_{12}} C_{m_{12},m_3,m_{12}+m_3}^{j_{12},j_3,j}, \end{aligned} \tag{4.13}$$

for fixed values of m_1, m_2 and m_3, and where $m_{12} = m_1 + m_2$. Keep m_{12} fixed, multiply both sides of (4.13) by $C_{m_1,m_2,m_{12}}^{j_1,j_2,j_{12}'}$, sum over m_1 and m_2 with $m_1 + m_2 = m_{12}$, and use (3.32); this leads to :

$$\begin{aligned} \sum_{m_1,m_2} & C_{m_1,m_2,m_{12}}^{j_1,j_2,j_{12}} C_{m_2,m_3,m_2+m_3}^{j_2,j_3,j_{23}} C_{m_1,m_2+m_3,m_1+m_2+m_3}^{j_1,j_{23},j} \\ & = U_{j_3,j,j_{23}}^{j_1,j_2,j_{12}} C_{m_{12},m_3,m_{12}+m_3}^{j_{12},j_3,j}, \end{aligned} \tag{4.14}$$

or

$$\begin{aligned} & U_{j_3,j,j_{23}}^{j_1,j_2,j_{12}} = \\ & \sum_{\substack{m_1,m_2 \\ m_1+m_2=m_{12}}} C_{m_1,m_2,m_{12}}^{j_1,j_2,j_{12}} C_{m_2,m_3,m_2+m_3}^{j_2,j_3,j_{23}} C_{m_1,m_2+m_3,m_{12}+m_3}^{j_1,j_{23},j} / C_{m_{12},m_3,m_{12}+m_3}^{j_{12},j_3,j}. \end{aligned} \tag{4.15}$$

Herein, m_{12} and m_3 are arbitrary but fixed, and the sum is over all m_1 and m_2 such that $m_1 + m_2 = m_{12}$. Let us now make the following choice : $m_{12} = j_{12}$ and $m_3 = j - j_{12}$. Then

$$U^{j_1,j_2,j_{12}}_{j_3,j,j_{23}} = \sum_{m_1} C^{j_1,j_2,j_{12}}_{m_1,j_{12}-m_1,j_{12}} C^{j_2,j_3,j_{23}}_{j_{12}-m_1,j-j_{12},j-m_1} C^{j_1,j_{23},j}_{m_1,j-m_1,j} / C^{j_{12},j_3,j}_{j_{12},j-j_{12},j},$$

or, rewritten an a more appropriate form :

$$U^{a,b,c}_{d,e,f} = \sum_{\alpha} C^{a,b,c}_{\alpha,c-\alpha,c} C^{b,d,f}_{c-\alpha,e-c,e-\alpha} C^{a,f,e}_{\alpha,e-\alpha,e} / C^{c,d,e}_{c,e-c,e}. \tag{4.16}$$

Using the closed form expression (3.21) for three of the four Clebsch-Gordan coefficients appearing in (4.16), there comes

$$U^{a,b,c}_{d,e,f} = \tag{4.17}$$
$$\sqrt{\frac{(2c+1)(a+b-c)!(a+f-e)!(c-d+e)!}{(-a+b+c)!(a-b+c)!(a+b+c+1)!(-a+f+e)!(a-f+e)!}}$$
$$\times\sqrt{\frac{(c+d+e+1)!}{(a+f+e+1)!}} \sum_{\alpha} \sqrt{\frac{(b+c-\alpha)!(f+e-\alpha)!}{(b-c+\alpha)!(f-e+\alpha)!}} \frac{(a+\alpha)!}{(a-\alpha)!} C^{b,d,f}_{c-\alpha,e-c,e-\alpha}.$$

Next, take the form (3.17) for the remaining Clebsch-Gordan coefficient. There comes :

$$U^{a,b,c}_{d,e,f} = \tag{4.18}$$
$$\sqrt{\frac{(2c+1)(2f+1)(a+b-c)!(a+f-e)!}{(-a+b+c)!(a-b+c)!(a+b+c+1)!(-a+f+e)!(a-f+e)!}}$$
$$\times\sqrt{\frac{(c-d+e)!(c+d+e+1)!(d-e+c)!(b+d-f)!}{(d+e-c)!(-b+d+f)!(b-d+f)!(b+d+f+1)!(a+f+e+1)!}}$$
$$\times\sum_{\alpha,l} \frac{(-1)^{b-c+\alpha-l}(b+c-\alpha+l)!(d+f-c+\alpha-l)!(a+\alpha)!(f+e-\alpha)!}{l!(f-e+\alpha-l)!(b-c+\alpha-l)!(d-f+c-\alpha+l)!(a-\alpha)!}.$$

Consider the summation part in here, and substitute $\alpha - l$ by k; there comes

$$\sum_{\alpha,k} (-1)^{b-c+k} \frac{(b+c-k)!(d+f-c+k)!(a+\alpha)!(f+e-\alpha)!}{(\alpha-k)!(f-e+k)!(b-c+k)!(d-f+c-k)!(a-\alpha)!}.$$

Since by (2.6)

$$\sum_{\alpha} \frac{(a+\alpha)!(f+e-\alpha)!}{(\alpha-k)!(a-\alpha)!} = \frac{(a+k)!(f+e+a+1)!(f+e-a)!}{(a-k)!(f+e+k+1)!},$$

the summation part in (4.18) is equal to

$$\sum_{k} (-1)^{b-c+k} \frac{(b+c-k)!(d+f-c+k)!(a+k)!(f+e+a+1)!(f+e-a)!}{(f-e+k)!(b-c+k)!(d-f+c-k)!(a-k)!(f+e+k+1)!}.$$

Putting this back in (4.18), and replacing the summation index k by $a-k$, there finally comes the following single sum expression :

$$U_{d,e,f}^{a,b,c} = (-1)^{a+b-c}\sqrt{(2c+1)(2f+1)}\frac{\nabla(a,e,f)\nabla(d,e,c)}{\nabla(a,b,c)\nabla(b,d,f)} \tag{4.19}$$
$$\times \sum_k \frac{(-1)^k(b+c-a+k)!}{k!(f-e+a-k)!(b-c+a-k)!}$$
$$\times \frac{(d+f-c+a-k)!(2a-k)!}{(d-f+c-a+k)!(1+f+e+a-k)!},$$

where

$$\nabla(a,b,c) = \sqrt{(-a+b+c)!(a-b+c)!(a+b+c+1)!/(a+b-c)!}.$$

One can think of $U_{d,e,f}^{a,b,c}$ as a function of six arguments belonging to $\frac{1}{2}\mathbb{N}$, which is zero whenever one of the triples (a,b,c), (d,e,c), (d,b,f) and (a,e,f) is not a triad, and which is given by (4.19) if all four triples are triads. In (4.19), the sum is over all k-values such that the arguments of the factorials are nonnegative, i.e. k runs from $\max(0, a-d+f-c)$ up to $\min(a+b-c, a+f-e)$.

In order to write this single sum in terms of a hypergeometric series, one must make an extra assumption. Let us assume that $-a+d+c-f \geq 0$ (otherwise, one should put $k' = -a+d+c-f+k$ and rewrite the expression above as a sum over k'). Then it follows directly that

$$U_{d,e,f}^{a,b,c} = (-1)^{a+b-c}\sqrt{(2c+1)(2f+1)}\frac{\nabla(a,e,f)\nabla(d,e,c)}{\nabla(a,b,c)\nabla(b,d,f)} \tag{4.20}$$
$$\times \frac{(b+c-a)!(d+f-c+a)!(2a)!}{(f-e+a)!(b-c+a)!(d-f+c-a)!(1+f+e+a)!}$$
$$\times\ {}_4F_3\left(\begin{matrix} 1-a+b+c, -a+e-f, -a-b+c, -1-a-e-f \\ -a+c-d-f, -2a, 1+d-f+c-a \end{matrix}; 1\right).$$

Many other expressions can be deduced by applying transformation (2.10) in a number of ways. For example, associating the $-n$ of (2.10) with $e-a-f$, and keeping the parameters $1-a+b+c$ and $-a+c-d-f$ fixed, the application of (2.10) leads to :

$${}_4F_3\left(\begin{matrix} 1-a+b+c, -a+e-f, -a-b+c, -1-a-e-f \\ -a+c-d-f, -2a, 1+d-f+c-a \end{matrix}; 1\right)$$
$$= \frac{(-a-b-c-1)_{a-e+f}(-b+d-f)_{a-e+f}}{(-2a)_{a-e+f}(1+d-f+c-a)_{a-e+f}}$$
$$\times\ {}_4F_3\left(\begin{matrix} -a+e-f, b-d-f, 1+c-d+e, 1-a+b+c \\ -a+c-d-f, -a+b-d+e+1, b+c+e-f+2 \end{matrix}; 1\right).$$

The Pochhammer symbols can be rewritten in terms of factorials. The actual expression depends on the sign of $(b+e-a-d)$; when $(b+e-a-d) \geq 0$, this leads to the following form :

$$
\begin{aligned}
U^{a,b,c}_{d,e,f} &= (-1)^{b+e-c-f}\sqrt{(2c+1)(2f+1)}\frac{\Delta(a,e,f)\Delta(b,d,f)}{\Delta(a,b,c)\Delta(c,d,e)} \qquad (4.21)\\
&\times\frac{(-a+b+c)!(c-d+e)!(a-c+d+f)!}{(a-e+f)!(-b+d+f)!(b+e-a-d)!(1+b+c+e-f)!}\\
&\times {}_4F_3\left(\begin{matrix}-a+e-f, b-d-f, 1+c-d+e, 1-a+b+c\\ -a+c-d-f, -a+b-d+e+1, b+c+e-f+2\end{matrix};1\right),
\end{aligned}
$$

where $\Delta(a,b,c)$ is given by (3.20).

Let us deduce one more expression, that is of interest for its symmetry. Apply (2.10) to the ${}_4F_3$ in (4.21), associating the $-n$ of (2.10) with $e-a-f$ and keeping the parameters $b-d-f$ and $1-a+b-d+e$ fixed; one finds that the ${}_4F_3$ of (4.21) is equal to

$$
\begin{aligned}
&\frac{(c-a-b)_{a+f-e}(2+c+d+e)_{a+f-e}}{(c-a-d-f)_{a+f-e}(2+b+c+e-f)_{a+f-e}}\\
&\times\ {}_4F_3\left(\begin{matrix}-a+e-f, b-d-f, b-a-c, e-c-d\\ -a+b-d+e+1, -a-c-d-f-1, b-c+e-f+1\end{matrix};1\right);
\end{aligned}
$$

assuming that $b+e-c-f\geq 0$, all Pochhammer symbols can be rewritten in terms of factorials, and putting this back in (4.21), one finds :

$$
\begin{aligned}
U^{a,b,c}_{d,e,f} &= \frac{(-1)^{b+e-c-f}\sqrt{(2c+1)(2f+1)}\Delta(a,b,c)\Delta(c,d,e)}{(a-b+c)!(c+d-e)!} \qquad (4.22)\\
&\times\frac{\Delta(a,e,f)\Delta(b,d,f)(1+a+c+d+f)!}{(a-e+f)!(-b+d+f)!(b+e-a-d)!(b+e-c-f)!}\\
&\times {}_4F_3\left(\begin{matrix}-a+e-f, b-d-f, b-a-c, e-c-d\\ -a+b-d+e+1, -a-c-d-f-1, b-c+e-f+1\end{matrix};1\right).
\end{aligned}
$$

So this expression is valid for $b+e\geq a+d$ and $b+e\geq c+f$ (which can always be obtained after applying a symmetry corresponding to a permutation of columns of the $6j$-coefficient). Rewriting with ${}_4F_3$ back as a single sum, one gets :

$$
\begin{aligned}
&U^{a,b,c}_{d,e,f} = \qquad (4.23)\\
&(-1)^{b+e-c-f}\sqrt{(2c+1)(2f+1)}\Delta(a,b,c)\Delta(c,d,e)\Delta(a,e,f)\Delta(b,d,f)\\
&\times\textstyle\sum_k(-1)^k(1+a+c+d+f-k)!/\big(k!(a-b+c-k)!(c+d-e-k)!\\
&\times(a-e+f-k)!(-b+d+f-k)!(b+e-a-d+k)!(b+e-c-f+k)!\big),
\end{aligned}
$$

or, when rewriting $a+c+d+f-k$ by k, one finds Racah's most symmetric formula

$$
\begin{aligned}
&U^{a,b,c}_{d,e,f} = \qquad (4.24)\\
&(-1)^{a+d+b+e}\sqrt{(2c+1)(2f+1)}\Delta(a,b,c)\Delta(c,d,e)\Delta(a,e,f)\Delta(b,d,f)\\
&\times \sum_k \frac{(-1)^k(1+k)!}{(k-a-b-c)!(k-c-d-e)!(k-d-b-f)!(k-a-e-f)!}\\
&\times \frac{1}{(a+d+b+e-k)!(a+d+c+f-k)!(b+e+c+f-k)!}.
\end{aligned}
$$

In (4.23) or (4.24), the sum is over all integer k-values such that all factorials assume nonnegative arguments.

4.4 The Racah polynomial

We are now in a position to describe the relation between $\mathfrak{su}(2)$ Racah coefficients and the so-called Racah polynomials. Consider expression (4.20), and let us rewrite :

$$
\begin{aligned}
&n = a-e+f,\ x = a+b-c,\ \alpha \equiv -N-1 = -a+c-d-f-1,\\
&\beta = -a+d-c-f-1,\ \gamma = -2a-1,\ \delta = 2c+1.
\end{aligned}
$$

Think of a, c, d and f as being fixed numbers (parameters), with

$$
c-a \geq |d-f| \qquad \text{and} \qquad c-d \geq |a-f|. \tag{4.25}
$$

Then b and e can be thought of as variables, with b varying between $c-a$ and $d+f$, and e running from $c-d$ to $a+f$. In terms of the new parameters/variables, this means that N is a fixed nonnegative integer parameter, and x and n are nonnegative integer variables with $0 \leq x \leq N$ and $0 \leq n \leq N$. The ${}_4F_3$-series of (4.20) is of the following form :

$$
\begin{aligned}
R_n(\lambda(x)) \equiv & R_n(\lambda(x);\alpha,\beta,\gamma,\delta) = \\
& {}_4F_3\left(\begin{matrix} -n, n+\alpha+\beta+1, -x, x+\gamma+\delta+1 \\ \alpha+1, \beta+\delta+1, \gamma+1 \end{matrix}; 1\right).
\end{aligned} \tag{4.26}
$$

In the case considered here, one of the denominator parameters ($\beta+\delta+1$) is strictly positive; the other two denominator parameters ($\alpha+1$ and $\gamma+1$) are negative integers with $\alpha+1=-N$ and $\gamma+1 \leq -N$.

In general, for given parameters α, β, γ, δ and any integer n, the expression (4.26) is a polynomial of degree n in the variable $\lambda(x) = x(x+\gamma+\delta+1)$. When one of the numerator parameters $\alpha+1$, $\beta+\delta+1$ or $\gamma+1$ is equal to a negative integer $-N$, then this polynomial is called a *Racah polynomial.* The other two numerator parameters should not belong to $\{0,-1,\ldots,-N+1\}$ in order to have (4.26) well defined for all integer n with $0 \leq n \leq N$.

In terms of the variables just introduced, one finds from (4.20) :

$$
\begin{aligned}
&U^{a,b,c}_{d,e,f} = \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (4.27)\\
&(-1)^x \frac{N!(-\gamma-1)!}{(\beta+\delta)!}\sqrt{\frac{(N+\gamma-\beta-n)!(-\beta-n-1)!}{n!(-\gamma-n-1)!(N-\beta-n)!(N-n)!}}\\
&\times\sqrt{\frac{\delta(N+1+\gamma-\beta)\,(\beta+\delta+n)!(N+\delta-n)!(\gamma+\delta+x)!(\beta+\delta+x)!}{x!(-\gamma-x-1)!(\delta+x)!(N-x)!(\gamma-\beta+x)!(N+1+\gamma+\delta+x)!}}\\
&\times R_n(\lambda(x);\alpha,\beta,\gamma,\delta).
\end{aligned}
$$

From (4.7) and (4.11), one has the following orthogonality :

$$
\sum_b (-1)^{e-e'}\frac{(2b+1)\sqrt{(2e+1)(2e'+1)}}{(2c+1)(2f+1)} U^{a,b,c}_{d,e,f} U^{a,b,c}_{d,e',f} = \delta_{e,e'}.
$$

In terms of the newly introduced parameters/variables, this can be rewritten as follows :

$$
\begin{aligned}
&\sum_{x=0}^{N} \frac{(\gamma+\delta+1+2x)(\gamma+\delta+1,\alpha+1,\beta+\delta+1,\gamma+1)_x}{(\gamma+\delta+1)x!(\gamma+\delta-\alpha+1,\gamma-\beta+1,\delta+1)_x} R_n(\lambda(x))R_{n'}(\lambda(x)) =\\
&\frac{(\gamma+\delta+2,-\beta)_N}{(\gamma-\beta+1,\delta+1)_N}\,\frac{n!(n+\alpha+\beta+1,\beta+1,\alpha-\delta+1,\alpha+\beta-\gamma+1)_n}{(\alpha+\beta+2)_{2n}(\alpha+1,\beta+\delta+1,\gamma+1)_n}\delta_{n,n'}.
\end{aligned}
$$

This is the orthogonality relation for Racah polynomials in the case that $N=-\alpha-1$; then $-\beta-\delta-1$ and $-\gamma-1$ should not belong to $\{0,1,\ldots,N-1\}$. Since the left hand side and right hand side of the last equation are rational functions of α, β, δ, γ, it follows that this equation is valid for all (real or complex) numbers such that both sides of the identity are defined. For given parameters, it is easy to check whether the weight function is positive, but the complete conditions for positivity are tedious to write down.

More generally, any of the three denominator parameters in (4.26) can stand for the negative integer $-N$. Then, one has :

$$
\begin{aligned}
&\sum_{x=0}^{N} \frac{(\gamma+\delta+1+2x)(\gamma+\delta+1)_x(\alpha+1)_x(\beta+\delta+1)_x(\gamma+1)_x}{(\gamma+\delta+1)x!(\gamma+\delta-\alpha+1)_x(\gamma-\beta+1)_x(\delta+1)_x}\\
&\qquad\qquad\times R_n(\lambda(x))R_{n'}(\lambda(x)) \qquad\qquad\qquad\qquad\qquad\qquad\qquad (4.28)\\
&= M\frac{n!(n+\alpha+\beta+1)_n(\beta+1)_n(\alpha-\delta+1)_n(\alpha+\beta-\gamma+1)_n}{(\alpha+\beta+2)_{2n}(\alpha+1)_n(\beta+\delta+1)_n(\gamma+1)_n}\delta_{n,n'},
\end{aligned}
$$

where

$$
M = \begin{cases}
\dfrac{(\gamma+\delta+2)_N(-\beta)_N}{(\gamma-\beta+1)_N(\delta+1)_N} & \text{if } \alpha+1=-N\\
\dfrac{(\gamma+\delta+2)_N(\delta-\alpha)_N}{(\gamma+\delta-\alpha+1)_N(\delta+1)_N} & \text{if } \beta+\delta+1=-N\\
\dfrac{(-\delta)_N(\alpha+\beta+2)_N}{(\alpha-\delta+1)_N(\beta+1)_N} & \text{if } \gamma+1=-N.
\end{cases}
$$

4.5 Wilson polynomials

Wilson polynomials can be seen as extensions of Racah polynomials. We will not prove the general form of the orthogonality relation for Wilson polynomials, but only indicate how it is related to that of Racah polynomials (we follow section 8.5.5 of [51] here).

Wilson realized that orthogonality still exists when the variable in the Racah polynomial is moved to the complex plane [54]. Thus he considered the more symmetric polynomials

$$\begin{aligned} p_n(t^2) \equiv p_n(t^2;a,b,c,d) &= (a+b)_n(a+c)_n(a+d)_n \\ &\times\ {}_4F_3\left({-n, a+b+c+d+n-1, a-t, a+t \atop a+b, a+c, a+d} ;1\right), \end{aligned} \tag{4.29}$$

having degree n in t^2. It follows immediately that these polynomials are symmetric under permutations of b, c and d. Furthermore, applying (2.10) to this ${}_4F_3$, keeping $-n$ and $a+b+c+d+n-1$ in the numerator and $a+b$ in the denominator fixed, one finds

$$p_n(t^2;a,b,c,d) = p_n(t^2;b,a,c,d).$$

Therefore, the Wilson polynomials are symmetric under permutations of a, b, c and d.

Following (4.26), one can write

$$p_n(t^2;a,b,c,d) = (a+b)_n(a+c)_n(a+d)_n\ R_n(\lambda(x);\alpha,\beta,\gamma,\delta),$$

where

$$\begin{aligned} a &= \gamma+\delta+\frac{1}{2}, \quad b=\alpha-\gamma+\delta-\frac{1}{2}, \\ c &= \beta-\gamma-\delta-\frac{1}{2}, \quad d=\gamma-\delta+\frac{1}{2}, \quad x=t-a. \end{aligned}$$

For $a+b=-N$, $N\in\mathbb{N}$, and $-a-c, -a-d \notin \{0,1,\ldots,N\}$, Wilson polynomials are proportional to Racah polynomials, so using (4.28) they satisfy the orthogonality relation

$$\begin{aligned} &\sum_{k=0}^{N} \frac{(2a)_k(a+1)_k(a+b)_k(a+c)_k(a+d)_k}{k!(a)_k(a-b+1)_k(a-c+1)_k(a-d+1)_k} \\ &\quad \times p_n((a+k)^2;a,b,c,d)p_m((a+k)^2;a,b,c,d) \\ &\quad = \frac{(2a+1)_N(1-c-d)_N}{(a-c+1)_N(a-d+1)_N} h_n\delta_{nm}, \end{aligned} \tag{4.30}$$

where $h_n =$

$$\frac{n!(a+b+c+d+n-1)_n(a+b)_n(a+c)_n(a+d)_n(b+c)_n(b+d)_n(c+d)_n}{(a+b+c+d)_{2n}}.$$

More generally, the orthogonality relation for the Wilson polynomials can be written in the following form [1], [54] :

$$\frac{1}{2\pi i}\int_{\mathcal{C}} p_n(z^2)p_m(z^2)\rho(z)dz = Mh_n\delta_{nm}, \tag{4.31}$$

where

$$M = \frac{2\Gamma(a+b)\Gamma(a+c)\Gamma(a+d)\Gamma(b+c)\Gamma(b+d)\Gamma(c+d)}{\Gamma(a+b+c+d)},$$
$$\rho(z) = \frac{\Gamma(a+z)\Gamma(a-z)\Gamma(b+z)\Gamma(b-z)\Gamma(c+z)\Gamma(c-z)\Gamma(d+z)\Gamma(d-z)}{\Gamma(2z)\Gamma(-2z)}.$$

The integration contour in (4.31) is the deformed imaginary axis which separates the poles

$$a+k, \quad b+k, \quad c+k, \quad d+k, \quad (k=0,1,2,\ldots)$$

and the poles

$$-a-k, \quad -b-k, \quad -c-k, \quad -d-k, \quad (k=0,1,2,\ldots)$$

of the function $\rho(z)$. Moreover, it is assumed that all the poles are distinct (so $2a$, $a+b$, $a+c$, $\ldots$, $c+d$, $2d$ are not in $\{-1,-2,\ldots\}$). When $a+b=-N$, $N\in\mathbb{N}$, the Wilson polynomials are proportional to Racah polynomials. One can deduce (4.30) from (4.31) by suitable integration contours and the Residue theorem [1], [54].

If a, b, c and d are positive real numbers, then one can take the imaginary axis as the contour $\mathcal{C}$ in (4.31). In that case, the orthogonality leads to [54]

$$\int_0^\infty p_n(-t^2)p_m(-t^2)w(t)dt = \delta_{nm}2\pi n!(a+b+c+d+n-1)_n\Gamma(n+a+b)$$
$$\times\frac{\Gamma(n+a+c)\Gamma(n+a+d)\Gamma(n+b+c)\Gamma(n+b+d)\Gamma(n+c+d)}{\Gamma(2n+a+b+c+d)},$$

where

$$w(t) = \left|\frac{\Gamma(a+it)\Gamma(b+it)\Gamma(c+it)\Gamma(d+it)}{\Gamma(2it)}\right|^2.$$

This can be further extended to :

Theorem 4.1. *Let a,b,c,d be complex parameters and*

$$W_n(x^2;a,b,c,d) = (a+b)_n(a+c)_n(a+d)_n \times {}_4F_3\left(\begin{matrix}-n, a+b+c+d+n-1, a-ix, a+ix\\ a+b, a+c, a+d\end{matrix};1\right) \tag{4.32}$$

be the Wilson polynomials. When $Re(a,b,c,d)>0$ and non-real parameters occur in conjugate pairs, Wilson polynomials are orthogonal on $\mathbb{R}_+$:

$$\begin{aligned}&\int_0^\infty \left|\frac{\Gamma(a+ix)\Gamma(b+ix)\Gamma(c+ix)\Gamma(d+ix)}{\Gamma(2ix)}\right|^2\\ &\quad\times W_n(x^2;a,b,c,d)W_m(x^2;a,b,c,d)\,dx\\ &=\delta_{nm}2\pi(n+a+b+c+d-1)_n n!\Gamma(n+a+b)\Gamma(n+a+c)\\ &\quad\times\frac{\Gamma(n+a+d)\Gamma(n+b+c)\Gamma(n+b+d)\Gamma(n+c+d)}{\Gamma(2n+a+b+c+d)}.\end{aligned} \tag{4.33}$$

For other cases where the measure has both an absolutely continuous and a discrete part, see [54], [25].

4.6 Racah coefficients of $\mathfrak{su}(1,1)$

Consider the tensor product of three irreducible $\star$-representations of $\mathfrak{su}(1,1)$, $\mathcal{D}_{k_1}\otimes\mathcal{D}_{k_2}\otimes\mathcal{D}_{k_3}$. Recall the decomposition of the tensor product of two such representations, given in (3.50). Following the same technique as in section 4.3, one can construct orthonormal basis vectors according to two different ways of "coupling", i.e.

$$(\mathcal{D}_{k_1}\otimes\mathcal{D}_{k_2})\otimes\mathcal{D}_{k_3} \quad\text{and}\quad \mathcal{D}_{k_1}\otimes(\mathcal{D}_{k_2}\otimes\mathcal{D}_{k_3}).$$

The corresponding basis vectors are determined by :

$$\begin{aligned}e_n^{((k_1k_2)k_{12}k_3)k} &= \sum_{n_{12},n_3} C^{k_{12},k_3,k}_{n_{12},n_3,n}\; e_{n_{12}}^{(k_1k_2)k_{12}}\otimes e_{n_3}^{(k_3)}\\ &= \sum_{n_1,n_2,n_3,n_{12}} C^{k_1,k_2,k_{12}}_{n_1,n_2,n_{12}} C^{k_{12},k_3,k}_{n_{12},n_3,n}\; e_{n_1}^{(k_1)}\otimes e_{n_2}^{(k_2)}\otimes e_{n_3}^{(k_3)},\end{aligned} \tag{4.34}$$

and

$$\begin{aligned}e_n^{(k_1(k_2k_3)k_{23})k} &= \sum_{n_1,n_{23}} C^{k_1,k_{23},k}_{n_1,n_{23},n}\; e_{n_1}^{(k_1)}\otimes e_{n_{23}}^{(k_2k_3)k_{23}}\\ &= \sum_{n_1,n_2,n_3,n_{23}} C^{k_2,k_3,k_{23}}_{n_2,n_3,n_{23}} C^{k_1,k_{23},k}_{n_1,n_{23},n}\; e_{n_1}^{(k_1)}\otimes e_{n_2}^{(k_2)}\otimes e_{n_3}^{(k_3)}.\end{aligned} \tag{4.35}$$

Following (3.51), the above representation labels are such that

$$\begin{aligned}&k_{12}=k_1+k_2+j_{12},\quad k_{23}=k_2+k_3+j_{23},\\ &k=k_{12}+k_3+j=k_1+k_{23}+j',\\ &j_{12},j,j_{23},j'\in\mathbb{Z}_+,\ \text{and}\ j_{12}+j=j_{23}+j'.\end{aligned} \tag{4.36}$$

As in (4.4) and (4.5), the Racah coefficient of $\mathfrak{su}(1,1)$ [for the positive discrete series representations] is then defined by

$$\langle e_n^{(k_1(k_2k_3)k_{23})k}, e_{n'}^{((k_1k_2)k_{12}k_3)k'}\rangle = \delta_{k,k'}\delta_{n,n'}U^{k_1,k_2,k_{12}}_{k_3,k,k_{23}}, \tag{4.37}$$

and therefore

$$e_n^{((k_1k_2)k_{12}k_3)k} = \sum_{k_{23}} U_{k_3,k,k_{23}}^{k_1,k_2,k_{12}} e_n^{(k_1(k_2k_3)k_{23})k}. \tag{4.38}$$

Herein, the sum is over all k_{23} such that (4.36) holds, i.e. over all integer j_{23} with $0 \leq j_{23} \leq j_{12} + j$ (where j_{12} and j are fixed by the left hand side of (4.38)). So the sum is finite. U is again an orthogonal matrix; with the convention (4.36) the orthogonality reads :

$$\sum_{k_{12}} U_{k_3,k,k_{23}}^{k_1,k_2,k_{12}} U_{k_3,k,k'_{23}}^{k_1,k_2,k_{12}} = \delta_{j_{23},j'_{23}}, \tag{4.39}$$

$$\sum_{k_{23}} U_{k_3,k,k_{23}}^{k_1,k_2,k_{12}} U_{k_3,k,k_{23}}^{k_1,k_2,k'_{12}} = \delta_{j_{12},j'_{12}}. \tag{4.40}$$

Following the technique of section 4.3, an explicit expression can be obtained for the Racah coefficient :

$$\begin{aligned} &U_{k_3,k,k_{23}}^{k_1,k_2,k_{12}} = \\ &\binom{j+j_{12}}{j_{23}} \frac{(2k_2)_{j_{12}}(2k_3)_j(2k_1+2k_2+2k_3+j+j_{12}-1)_{j_{23}}}{(2k_3,2k_2+2k_3+j_{23}-1)_{j_{23}}(2k_2+2k_3+2j_{23})_{j'}} \left(\frac{j'!\,j_{23}!}{j!\,j_{12}!}\right)^{1/2} \\ &\times \left(\frac{(2k_1,2k_{23},2k_1+2k_{23}+j'-1)_{j'}\ (2k_2,2k_3,2k_2+2k_3+j_{23}-1)_{j_{23}}}{(2k_{12},2k_3,2k_{12}+2k_3+j-1)_j\ (2k_1,2k_2,2k_1+2k_2+j_{12}-1)_{j_{12}}}\right)^{1/2} \\ &\times {}_4F_3\left({2k_1+2k_2+j_{12}-1, 2k_2+2k_3+j_{23}-1, -j_{12}, -j_{23} \atop 2k_2, 2k_1+2k_2+2k_3+j+j_{12}-1, -j-j_{12}}; 1\right), \end{aligned} \tag{4.41}$$

again with the labels determined by (4.36).

Orthogonality relations for $\mathfrak{su}(1,1)$ Racah coefficients are again equivalent with the orthogonality of Racah polynomials.

4.7 Notes and Exercises

Racah coefficients appeared for the first time in Racah's 1942 paper, dealing with calculations in many-electron spectra [33]. In that same paper, Racah derived already the single sum expression (4.23) in an appendix. The actual ${}_4F_3$ expressions for these coefficients were given only later; see [50] or [36] for a summary of such expressions and references. The notation and introduction of $6j$-coefficients is due to Wigner [52]. He also gave the classical symmetries of $6j$-coefficients, such as (4.11). Further symmetries, closely related to the ${}_4F_3$ expression, were given by Regge [38].

The relation between $\mathfrak{su}(2)$ Racah coefficients and orthogonal polynomials (from then on referred to as Racah polynomials) was discovered only much later [2]. It led also to the discovery of Wilson polynomials [54] and Askey-Wilson polynomials [3].

Racah coefficients and $6j$-coefficients play a crucial role in nuclear, atomic and molecular physics, through the importance of angular momentum properties. The collection of formulas of [50] is again very useful for people working in the field of "Racah-Wigner algebra" (see also [7], volume 9).

Exercises

1. Assume that (a,b,c) forms a triad. Verify that
$$\begin{Bmatrix} a & b & c \\ d & e & 0 \end{Bmatrix} = (-1)^{a+b+c} \frac{\delta_{a,e}\delta_{b,d}}{\sqrt{(2a+1)(2b+1)}}. \tag{4.42}$$
2. Rewrite the orthogonality relations for Racah coefficients in terms of $6j$-coefficients :
$$\sum_x (2x+1) \begin{Bmatrix} a & b & x \\ c & d & p \end{Bmatrix} \begin{Bmatrix} a & b & x \\ c & d & q \end{Bmatrix} = \frac{\delta_{p,q}}{2p+1}, \tag{4.43}$$
provided (a,d,p) and (b,c,p) form triads.
3. Assume that (a,b,c) forms a triad. Show that
$$\sum_x (-1)^{a+b+x}(2x+1) \begin{Bmatrix} a & b & x \\ b & a & c \end{Bmatrix} = \delta_{c,0}\sqrt{(2a+1)(2b+1)}, \tag{4.44}$$
where the sum is over all (integer or half-integer) x which obey all triangular conditions.

5 Tensor product of $n+1$ representations and $3nj$-coefficients

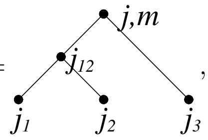
We shall consider the tensor product of $n+1$ irreducible $\star$-representations of $\mathfrak{su}(2)$. The definition of binary coupling schemes and the related $3nj$-coefficients is discussed. Some typical problems in this context are of combinatorial type; for $(n+1)$-fold tensor products of $\mathfrak{su}(1,1)$ these problems are exactly the same. Some other problems are related to realizations of basis vectors, and lead to multivariable orthogonal polynomials.

5.1 $3nj$-coefficients

Consider the following tensor product of $n+1$ $\mathfrak{su}(2)$ representations,
$$V = D_{j_1} \otimes D_{j_2} \otimes \cdots \otimes D_{j_{n+1}}.$$
A basis for V (the uncoupled basis) is given by
$$e^{(j_1)}_{m_1} \otimes e^{(j_2)}_{m_2} \otimes \cdots \otimes e^{(j_{n+1})}_{m_{n+1}},$$
where m_i is a projection of j_i.

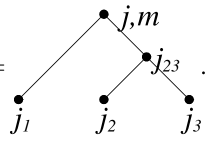
Coupled basis vectors can be defined by means of *binary coupling schemes*. The idea is a simple extension of the two ways in which the tensor product of

three representations can be "coupled", see (4.1). The vector (4.2) is depicted by the following coupling scheme :

$$e_m^{((j_1 j_2) j_{12} j_3) j} = T_1(j,m) =$$ 

,

and the vector (4.3) by :

$$e_m^{(j_1 (j_2 j_3) j_{23}) j} = T_2(j,m) =$$ 

.

The Racah coefficient (or, loosely speaking, the $6j$-coefficient) is then

$$U_{j_3,j,j_{23}}^{j_1,j_2,j_{12}} = \langle T_2(j,m), T_1(j,m)\rangle. \tag{5.1}$$

Generally, a binary coupling scheme $T = T(J,M)$ is defined recursively.

Definition 5.1. *A binary coupling scheme on a twofold tensor product $D_j \otimes D_{j'}$ is defined as*

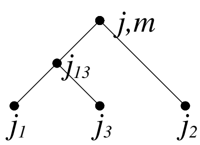
$$\sum_{m,m'} C_{m,m',M}^{j,j',J}\; e_m^{(j)} \otimes e_{m'}^{(j')} \qquad \textit{or} \qquad \sum_{m,m'} C_{m',m,M}^{j',j,J}\; e_m^{(j)} \otimes e_{m'}^{(j)'}.$$

A binary coupling scheme $T = T(J,M)$ on the tensor product of $(n+1)$ representations $D_{j_1} \otimes D_{j_2} \otimes \cdots \otimes D_{j_{n+1}}$ is defined as

$$\sum_{m,m'} C_{m,m',M}^{j,j',J}\; \sigma(T_1(j,m) \otimes T_2(j',m')),$$

where $T_1(j,m)$ is a binary coupling scheme on the tensor product of l representations and $T_2(j',m')$ is a binary coupling scheme on the tensor product of the remaining $n+1-l$ representations. The map σ reshuffles the order of the components in the tensor product such that they belong to $D_{j_1} \otimes D_{j_2} \otimes \cdots \otimes D_{j_{n+1}}$.

For example, in $D_{j_1} \otimes D_{j_2} \otimes D_{j_3}$:

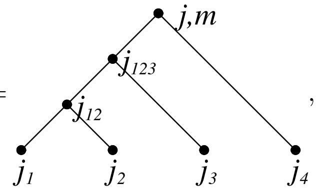


$$= e_m^{((j_1 j_3) j_{13} j_3) j} = \sum_{m_{13},m_2} C_{m_{13},m_2,m}^{j_{13},j_2,j}\; \sigma(e_{m_{13}}^{(j_1 j_3) j_{13}} \otimes e_{m_2}^{(j_2)})$$

$$= \sum_{\substack{m_1,m_2,m_3\\ m_1+m_2+m_3=m}} C_{m_1,m_3,m_{13}}^{j_1,j_3,j_{13}} C_{m_{13},m_2,m}^{j_{13},j_2,j}\; e_{m_1}^{(j_1)} \otimes e_{m_2}^{(j_2)} \otimes e_{m_3}^{(j_3)}.$$

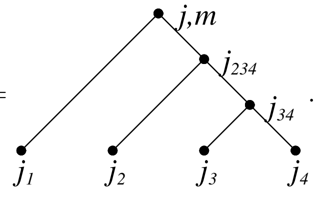
The following are two binary coupling schemes on $D_{j_1} \otimes D_{j_2} \otimes D_{j_3} \otimes D_{j_4}$ $(n = 3)$:

$$T_1 \equiv T_1(j,m) = e_m^{(((j_1 j_2)j_{12} j_3)j_{123} j_4)j} =$$



, (5.2)

$$T_2 \equiv T_2(j,m) = e_m^{(j_1(j_2(j_3 j_4)j_{34})j_{234})j} =$$



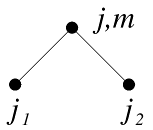
. (5.3)

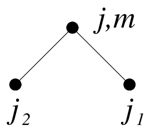
In a binary coupling scheme, we shall use the terminology of binary trees : for example, in the above scheme T_2, j_1, j_2, j_3 and j_4 are the labels of the *leaf nodes* of the tree, j_{34} and j_{234} are the labels of the *internal nodes*, and j (or j,m) is the label of the *top node*.

Definition 5.2. *A* $3nj$*-coefficient* ($n > 1$) *is a particular overlap coefficient* $\langle T_2, T_1\rangle = \langle T_2(J,M), T_1(J,M)\rangle$, *where* $T_1(J,M)$ *and* $T_2(J,M)$ *are two binary coupling schemes on the tensor product of* $(n+1)$ *representations* $D_{j_1} \otimes D_{j_2} \otimes \cdots \otimes D_{j_{n+1}}$.

Observe that, just as for a $6j$-coefficient, the $3nj$-coefficient is independent of the M-value.

For $n = 1$, there is essentially only one binary coupling scheme. In fact, there are two binary coupling schemes, namely

$e_m^{(j_1 j_2)j} =$ 

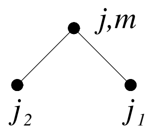
and $e_m^{(j_2 j_1)j} =$ 

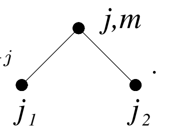
;

but since $C^{j_2,j_1,j}_{m_2,m_1,m} = (-1)^{j_1+j_2-j} C^{j_1,j_2,j}_{m_1,m_2,m}$, see (3.25),

$$e_m^{(j_2 j_1)j} = \sum_{m_1,m_2} C^{j_2,j_1,j}_{m_2,m_1,m} e^{(j_1)}_{m_1} \otimes e^{(j_2)}_{m_2} = (-1)^{j_1+j_2-j} e_m^{(j_1 j_2)j}.$$

This can be written as

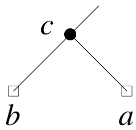


$= (-1)^{j_1+j_2-j}$


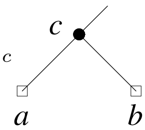
.

By definition of binary coupling schemes, it follows that this rule is also valid for arbitrary binary coupling schemes :

Rule 1 :



$= (-1)^{a+b-c}$

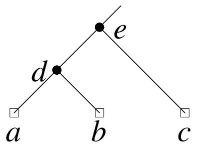


, (5.4)

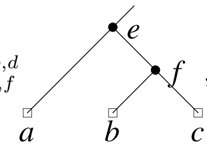
where a and b can be representation labels of leaf nodes or internal nodes, and c is the representation label of an internal node or the top node.

A second rule that can be used on binary coupling schemes is the following :

Rule 2 :



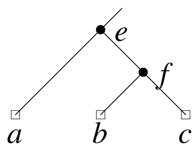
$= \sum_f U^{a,b,d}_{c,e,f}$



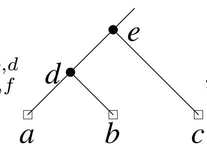
, (5.5)

or equivalently (using Rule 1 and symmetries of U) :

Rule 2′ :



$= \sum_d U^{a,b,d}_{c,e,f}$



, (5.6)

Herein, a, b and c are labels of leaf nodes or internal nodes; d and f are labels of internal nodes; and e is the label of an internal node or the top node. This follows from the general definition, and (5.1).

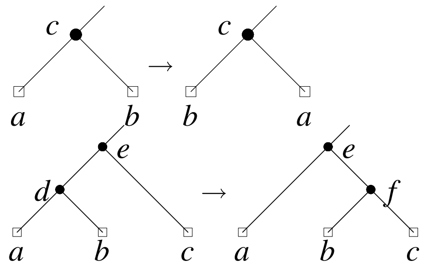
All the notions introduced in this subsection for $\mathfrak{su}(2)$ are also appropriate for the case of tensor products of positive discrete series representations of $\mathfrak{su}(1,1)$. In particular, Definition 5.1 remains valid up to the replacement of C by the $\mathfrak{su}(1,1)$ Clebsch-Gordan coefficient (3.52). Also Definition 5.2 remains valid up to the trivial replacements of $\mathfrak{su}(2)$ labels (j,m) by $\mathfrak{su}(1,1)$ labels (k,n). Moreover, since the $\mathfrak{su}(1,1)$ Clebsch-Gordan coefficients (3.52) satisfy the same symmetry (3.55), Rule 1 remains valid for $\mathfrak{su}(1,1)$ binary coupling schemes. By (4.38) it follows that also Rule 2 is valid for $\mathfrak{su}(1,1)$ binary coupling schemes, where U is an $\mathfrak{su}(1,1)$ Racah coefficient (4.41). So as far as binary coupling schemes are concerned, we can work with the two rules (5.4) and (5.5) both in the $\mathfrak{su}(2)$ and in the $\mathfrak{su}(1,1)$ case.

5.2 An example : the Biedenharn-Elliott identity

Consider two binary coupling schemes T_1 and T_2. Our purpose is to compute the $3nj$-coefficient $\langle T_2, T_1\rangle$ in terms of Racah coefficients. This method of computation is often referred to as the "method of trees". By the remark in section 5.1, this method works both in the $\mathfrak{su}(2)$ and the $\mathfrak{su}(1,1)$ case. For the computation of $\langle T_2, T_1\rangle$, one can use the two rules

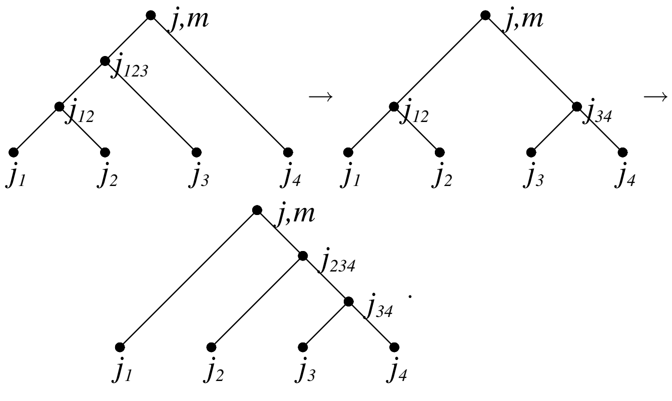
Rule 1 : $\rightarrow$ $(-1)^{a+b-c}$

Rule 2 : $\rightarrow$ $\sum_f U^{a,b,d}_{c,e,f}$ (5.7)



to go from T_1 to T_2. The last column gives the effect of applying the rule. A consecutive sequence of applications of Rule 1 and Rule 2 to go from one binary coupling scheme to another one, is called a *path* from T_1 to T_2.

Consider now the two binary coupling schemes T_1 and T_2 defined in (5.2) and (5.3). One path from T_1 to T_2 is given by



(5.8)

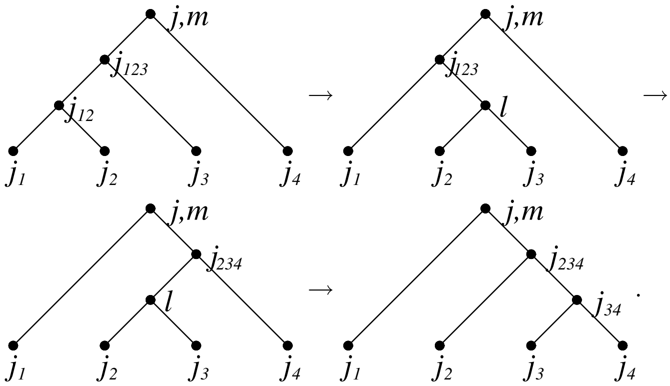
This means

$$T_1 = e_m^{(((j_1j_2)j_{12}j_3)j_{123}j_4)j} = \sum_{j'_{34}} U^{j_{12},j_3,j_{123}}_{j_4,j,j'_{34}} e_m^{((j_1j_2)j_{12}(j_3j_4)j'_{34})j}$$
$$= \sum_{j'_{34},j'_{234}} U^{j_{12},j_3,j_{123}}_{j_4,j,j'_{34}} U^{j_1,j_2,j_{12}}_{j_{34},j,j'_{234}} e_m^{(j_1(j_2(j_3j_4)j'_{34})j'_{234})j},$$

or

$$\langle T_2, T_1\rangle = U^{j_{12},j_3,j_{123}}_{j_4,j,j_{34}} U^{j_1,j_2,j_{12}}_{j_{34},j,j_{234}}. \tag{5.9}$$

Another path from T_1 to T_2 is given by



(5.10)

This means

$$\begin{aligned} T_1 = e_m^{(((j_1j_2)j_{12}j_3)j_{123}j_4)j} &= \sum_l U^{j_1,j_2,j_{12}}_{j_3,j_{123},l} e_m^{(((j_1(j_2j_3)l)j_{123}j_4)j} \\ &= \sum_{l,j'_{234}} U^{j_1,j_2,j_{12}}_{j_3,j_{123},l} U^{j_1,l,j_{123}}_{j_4,j,j'_{234}} e_m^{(j_1((j_2j_3)lj_4)j'_{234})j} \\ &= \sum_{l,j'_{234},j'_{34}} U^{j_1,j_2,j_{12}}_{j_3,j_{123},l} U^{j_1,l,j_{123}}_{j_4,j,j'_{234}} U^{j_2,j_3,l}_{j_4,j'_{234},j'_{34}} e_m^{(j_1(j_2(j_3j_4)j'_{34})j'_{234})j}, \end{aligned}$$

or

$$\langle T_2, T_1\rangle = \sum_l U^{j_1,j_2,j_{12}}_{j_3,j_{123},l} U^{j_1,l,j_{123}}_{j_4,j,j_{234}} U^{j_2,j_3,l}_{j_4,j_{234},j_{34}}. \qquad (5.11)$$

Comparison of (5.9) and (5.11) leads to :

Theorem 5.1. *The Racah coefficients of $\mathfrak{su}(2)$ satisfy the following identity, known as the Biedenharn-Elliott identity :*

$$U^{j_1,j_2,j_{12}}_{j_{34},j,j_{234}} U^{j_{12},j_3,j_{123}}_{j_4,j,j_{34}} = \sum_l U^{j_1,j_2,j_{12}}_{j_3,j_{123},l} U^{j_1,l,j_{123}}_{j_4,j,j_{234}} U^{j_2,j_3,l}_{j_4,j_{234},j_{34}}. \qquad (5.12)$$

In terms of $6j$-coefficients, this can be rewritten as :

$$\sum_x (-1)^{R+x}(2x+1) \begin{Bmatrix} a & b & x \\ c & d & p \end{Bmatrix} \begin{Bmatrix} c & d & x \\ e & f & q \end{Bmatrix} \begin{Bmatrix} e & f & x \\ b & a & r \end{Bmatrix} = \qquad (5.13)$$
$$\begin{Bmatrix} p & q & r \\ e & a & d \end{Bmatrix} \begin{Bmatrix} p & q & r \\ f & b & c \end{Bmatrix},$$

where $R = a+b+c+d+e+f+p+q+r$, and all labels are representation labels (elements of $\frac{1}{2}\mathbb{N}$). The sum is over all x with $\max(|a-b|, |c-d|, |e-f|) \leq x \leq \min(a+b, c+d, e+f)$.

The Racah coefficients of $\mathfrak{su}(1,1)$ satisfy the same identity (Biedenharn-Elliott identity for $\mathfrak{su}(1,1)$) :

$$U^{k_1,k_2,k_{12}}_{k_{34},k,k_{234}} U^{k_{12},k_3,k_{123}}_{k_4,k,k_{34}} = \sum_K U^{k_1,k_2,k_{12}}_{k_3,k_{123},K} U^{k_1,K,k_{123}}_{k_4,k,k_{234}} U^{k_2,k_3,K}_{k_4,k_{234},k_{34}}. \tag{5.14}$$

All labels are $\mathfrak{su}(1,1)$ representation labels, and the sum over K is finite with K running from k_2+k_3 to $\min(k_{123}-k_1, k_{234}-k_4)$ in steps of 1.

5.3 A recurrence relation for Racah polynomials

We shall briefly indicate how the Biedenharn-Elliott identity can be used to derive the recurrence relation for Racah polynomials. Consider (5.13) with $q=c$, $r=b$ and $f=1$, then

$$\sum_x (-1)^{R+x}(2x+1) \begin{Bmatrix} a & b & x \\ c & d & p \end{Bmatrix} \begin{Bmatrix} c & d & x \\ e & 1 & c \end{Bmatrix} \begin{Bmatrix} e & 1 & x \\ b & a & b \end{Bmatrix} =$$
$$\begin{Bmatrix} p & c & b \\ e & a & d \end{Bmatrix} \begin{Bmatrix} p & c & b \\ 1 & b & c \end{Bmatrix},$$

where the sum is over $x \in \{e-1, e, e+1\}$; or rewritten in terms of Racah coefficients,

$$\sum_{l=-1}^{1} U^{a,d,p}_{c,b,e+l} U^{a,e+l,b}_{1,b,e} U^{c,d,e+l}_{1,e,c} = U^{a,d,p}_{c,b,e} U^{p,c,b}_{1,b,c}. \tag{5.15}$$

All the Racah coefficients with a representation label 1 as one of their parameters simplify; e.g. using (4.24) the summation $\sum_k$ is seen to reduce to only two terms, so each such coefficient consists of certain factors times two terms. Next, apply the symmetry (4.11) to the remaining Racah coefficients in (5.15). The result is a linear combination of the following form :

$$g_1 U^{a,p,d}_{c,e-1,b} + g_2 U^{a,p,d}_{c,e,b} + g_3 U^{a,p,d}_{c,e+1,b} = g_4 U^{a,p,d}_{c,e,b}. \tag{5.16}$$

Next, using the explicit expressions of these simple coefficients and the relation (4.27) between Racah coefficients and Racah polynomials, one finds :

$$A_n R_{n+1}(\lambda(x)) - (A_n + C_n) R_n(\lambda(x)) + C_n R_{n-1}(\lambda(x)) = \lambda(x) R_n(\lambda(x)),$$

where

$$R(\lambda(x)) = R_n(\lambda(x); \alpha, \beta, \gamma, \delta),$$

and

$$A_n = \frac{(n+\alpha+\beta+1)(n+\alpha+1)(n+\beta+\delta+1)(n+\gamma+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)},$$
$$C_n = \frac{n(n+\beta)(n+\alpha+\beta-\gamma)(n+\alpha-\delta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+1)}.$$

This is the three-term recurrence relation for Racah polynomials.

5.4 Expressions for $3nj$-coefficients

In section 5.2 we have computed $\langle T_2, T_1 \rangle$, with T_1 and T_2 given in (5.2) and (5.3), in two different ways : one expression is (5.9), the other is (5.11). It is obvious that from a computational point of view, (5.9) is much better than (5.11). Indeed, (5.9) is simply the product of two Racah coefficients, whereas (5.11) is a single sum over the product of three Racah coefficients. This is related to the length of the corresponding paths, given in (5.8) and (5.10), to go from T_1 to T_2.

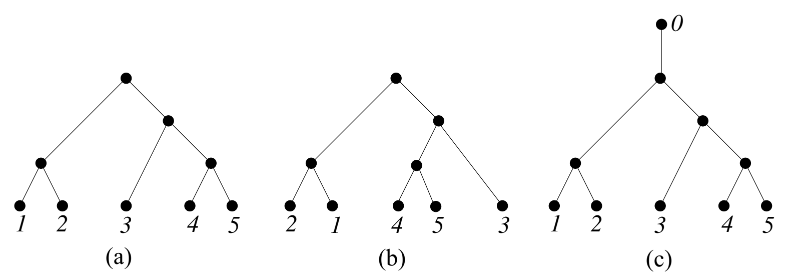
In general, consider two binary coupling schemes T_1 and T_2. In order to compute the $3nj$-coefficient $\langle T_2, T_1 \rangle$, one must find a path from T_1 to T_2 using the rules (5.7). Theoretically, every path will do; but in order to find an efficient way of computing the $3nj$-coefficient, one must find a shortest path, i.e. a path for which the number of applications of Rule 2 is minimal. Indeed, Rule 1 does not have any effect on the efficiency of the formula (it is only a $\pm$-sign). But Rule 2 introduces a Racah coefficient and a summation variable. Clearly, a formula with the least number of summation variables is the most optimal.

We can redefine the problem of finding an optimal expression for a $3nj$-coefficient by introducing *binary coupling trees.* A binary coupling tree on $n+1$ leaves is a rooted binary tree such that :

- the $n+1$ leaves are labelled $1, 2, \ldots, n+1$;
- the internal nodes are not labelled;
- for each internal node, one can exchange the left and right children of that node.

These are sometimes referred to as *unordered rooted binary trees with labelled leaves.* An example is given in Figure 1. Note that in this figure, (a) and (b)

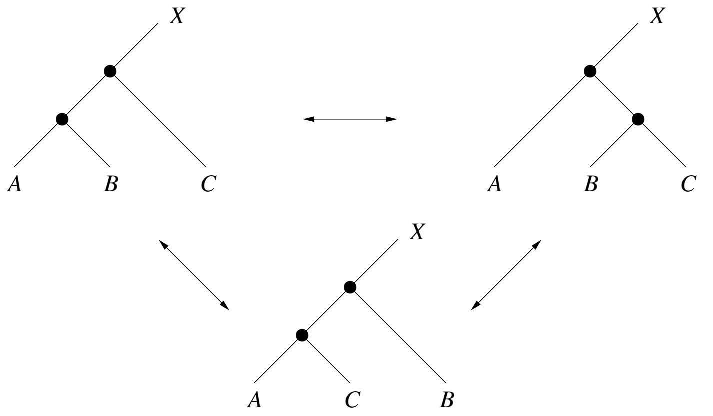
Fig. 1. Binary coupling trees



represent the same binary coupling tree, since one can freely exchange the left and right children. Sometimes it will be convenient to attach an extra node with label 0 to the root of the binary coupling tree, such as in (c). The only

elementary operation that now remains is the one corresponding to Rule 2; this is referred to as *a rotation* on binary coupling trees (since by definition the operation corresponding to Rule 1 does not change the binary coupling tree). This is illustrated in Figure 2, where A, B and C represent subtrees and X is a part of the binary coupling tree containing the root (or, equivalently, the label 0).

Fig. 2. Binary coupling trees related by a rotation



The relation with binary coupling schemes is obvious. The leaf labels $1, 2, \ldots, n+1$ refer to the representation labels $j_1, j_2, \ldots, j_{n+1}$. An internal node is no longer explicitly labelled, but it is implicitly labelled by the collection of leaves underneath it. For a given $3nj$-coefficient with binary coupling schemes T_1 and T_2 and for every path from T_1 to T_2 there exists a sequence of rotations transforming the corresponding binary coupling trees into each other and vice versa. Clearly, from the sequence of rotations between the binary coupling trees, the path between the binary coupling schemes can be reconstructed and hence no information is lost for determining the expression for the $3nj$-coefficient. The problem is thus reduced to finding a shortest sequence of rotations transforming one binary coupling tree into another.

A binary coupling tree can be given either explicitly as a graph, see Figure 1, or as a bracketing of the leaf labels. For example, the binary coupling tree of Figure 1 could be represented as

$$((1,2),(3,(4,5))) \qquad \text{or} \qquad ((2,1),((4,5),3)). \tag{5.17}$$

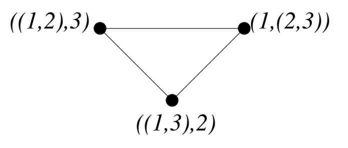
Henceforth we shall use this notation for a binary coupling tree.

Let $n > 1$ be fixed, and consider the set of all binary coupling trees with $n+1$ leaves. Since our only basic operation is rotation, we shall consider the *rotation graph* of binary coupling trees. This graph G_n has as vertex set

the set of all binary coupling trees with $n+1$ leaves, and there is an edge between two vertices if and only if the corresponding binary coupling trees are related through a single rotation. It follows that an optimal expression for a $3nj$-coefficient corresponds to finding a shortest path in G_n between the two binary coupling trees corresponding to the two binary coupling schemes in the $3nj$-coefficient.

We shall now consider some examples, and deduce some general properties of G_n. For $n=2$ this graph is simply a triangle. In Figure 3 we give G_2, and use the convention (5.17) to label the corresponding binary coupling trees. The next graph, G_3, has order 15. This graph is shown in Figure 4, using the

Fig. 3. The graph G_2



bracket representation (5.17) for the binary coupling trees. Observe that the two paths described in (5.8) and (5.10) correspond to two paths in G_3 : (5.8) corresponds to $(((1,2),3),4) \rightarrow ((1,2),(3,4)) \rightarrow (1,(2,(3,4)))$; (5.10) corresponds to $(((1,2),3),4) \rightarrow (((1,(2,3)),4) \rightarrow (1,((2,3),4)) \rightarrow (1,(2,(3,4)))$.

Let us now consider some general properties of the graph G_n. An arbitrary element of G_n, i.e. a binary coupling tree on $n+1$ leaves, has $n-1$ internal edges (i.e. edges containing no leaf). Two rotations can be performed with respect to each internal edge, thus every binary coupling tree is connected by an edge to $2(n-1)$ other binary coupling trees. In other words, G_n is a regular graph of degree $2(n-1)$. For example, G_3 is regular of degree 4.

To determine the number of binary coupling trees on $n+1$ leaves (or the order $|G_n|$ of G_n), consider first a binary coupling tree on n leaves (with labels $1,\ldots,n$), and extend the root of this tree with an extra edge ending in the leaf 0 (as in Figure 1). This tree has $2n-1$ edges in total. Therefore, there are $2n-1$ different ways of adding an extra edge ending with leaf label $n+1$ to this tree, namely by attaching it to each consisting edge, see e.g. Figure 5. Thus we have $|G_n| = (2n-1)|G_{n-1}|$, and find

$$|G_n| = (2n-1)!! = (2n-1)(2n-3)\cdots 3\cdot 1. \tag{5.18}$$

This implies that the order of G_n grows exponentially.

It is also easy to show that G_n is a connected graph, i.e. for any two binary coupling trees there exists at least one path between them.

The problem of determining the distance between two binary coupling trees of G_n (and thus of finding an optimal expression for the corresponding $3nj$-coefficient) turns out to be very hard. In [15], [16], distance properties in

Fig. 4. The graph G_3

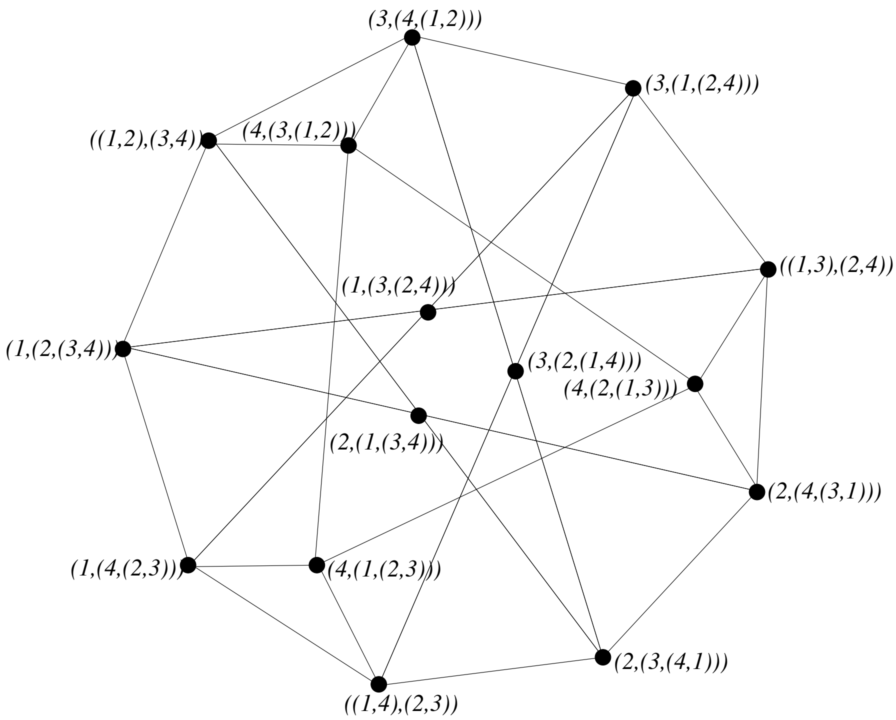
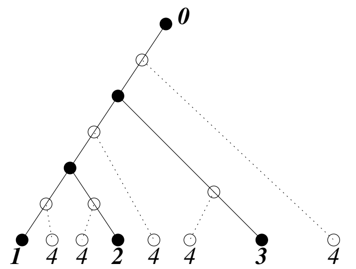


Fig. 5. Five ways of attaching an extra leaf label 4 to a given binary coupling tree on labels 1,2,3



G_n were studied. In particular, bounds for the *diameter* of G_n (this is the maximal length of all shortest paths) were found :

Theorem 5.2. *The diameter* $d(G_n)$ *of* G_n *satisfies*

$$\frac{1}{4}\log(n!) < d(G_n) < n\lceil\log(n)\rceil + n - 2\lceil\log(n)\rceil + 1. \tag{5.19}$$

Herein, $\lceil x\rceil$ *is the smallest integer larger than or equal to* $x > 0$*, and* $\log = \log_2$.

This implies that in an optimal expression of a $3nj$-coefficient in terms of Racah coefficients, the number of Racah coefficients appearing in a term of the expansion is of order $n\log(n)$.

All the considerations of this subsection are applicable both to $3nj$-coefficients of $\mathfrak{su}(2)$ and $\mathfrak{su}(1,1)$.

5.5 Multivariable Jacobi polynomials associated with binary coupling trees

Recall the definition of the Racah polynomial (4.26) and the Wilson polynomial (4.32). Other polynomials of hypergeometric type can be obtained as limits of Racah or Wilson polynomials [25].

For example, the continuous Hahn polynomials, denoted $p_m(x;a,b,c,d)$, are defined as [25]

$$\begin{aligned} p_m(x;a,b,c,d) =& i^m \frac{(a+c)_m(a+d)_m}{m!} \\ & \times {}_3F_2\left({-m, m+a+b+c+d-1, a+ix \atop a+c, a+d}; 1\right); \end{aligned} \tag{5.20}$$

for their orthogonality (when $\mathrm{Re}(a,b,c,d) > 0$, $\bar{c} = a$ and $\bar{d} = b$), see [25]. They can be obtained from the Wilson polynomials using the limit transition [25] :

$$\lim_{t\to-\infty} \frac{W_m((x-t)^2; a+it, b+it, c-it, d-it)}{(2t)^m m!} = p_m(x;a,b,c,d). \tag{5.21}$$

The classical Jacobi polynomials are defined by :

$$P_m^{(\alpha,\beta)}(x) = \frac{(\alpha+1)_m}{m!} {}_2F_1\left({-m, m+\alpha+\beta+1 \atop \alpha+1}; \frac{1-x}{2}\right); \tag{5.22}$$

for $\alpha, \beta > -1$, they are orthogonal over the interval $[-1,1]$ for the weight function $(1-x)^\alpha(1+x)^\beta$.

We shall now deduce some properties of orthogonal polynomials directly from the Biedenharn-Elliott identity for $\mathfrak{su}(1,1)$, see (5.14),

$$U^{k_1,k_2,k_{12}}_{k_{34},k,k_{234}} U^{k_{12},k_3,k_{123}}_{k_4,k,k_{34}} = \sum_K U^{k_1,k_2,k_{12}}_{k_3,k_{123},K} U^{k_1,K,k_{123}}_{k_4,k,k_{234}} U^{k_2,k_3,K}_{k_4,k_{234},k_{34}}, \tag{5.23}$$

where U is an $\mathfrak{su}(1,1)$ Racah coefficient.

Theorem 5.3. *The Wilson polynomials satisfy the following convolution identity :*

$$\begin{aligned}
&\sum_{l=0}^{m+j} \binom{j+m}{l} \frac{(2k_2)_m(2k_3)_j(2k_1+2k_2+2k_3+j+m-1)_l}{(2k_3)_l(2k_2+2k_3+l-1)_l(2k_2+2k_3+2l)_{j+m-l}} \\
&\times R_l(\lambda(m);2k_2-1,2k_3-1,-j-m-1,2k_1+2k_2+j+m-1) \\
&\times W_{m+j-l}((x_1-t)^2;k_1+it,k_2+k_3+l-is+it, \\
&\qquad\qquad k_1-it,k_2+k_3+l+is-it) \\
&\times W_l((x_1+x_2-t)^2;k_2-ix_1+it,k_3-is+it,k_2+ix_1-it,k_3+is-it) \\
&= W_m((x_1-t)^2;k_1+it,k_2-i(x_1+x_2)+it,k_1-it,k_2+i(x_1+x_2)-it) \\
&\times W_j((x_1+x_2-t)^2;k_1+k_2+m+it, \\
&\qquad\qquad k_3-is+it,k_1+k_2+m-it,k_3+is-it),
\end{aligned} \tag{5.24}$$

where $j,m\in\mathbb{N}$, $k_1,k_2,k_3,x_1,x_2,x_3,t\in\mathbb{R}$ *and* $s=x_1+x_2+x_3$.

Proof. In the case of $\mathfrak{su}(1,1)$, the summation range for K in (5.23) is from k_2+k_3 to $\min(k_{123}-k_1,k_{234}-k_4)$. The summation variable K thus takes real values, starting with k_2+k_3 and increasing in steps of one. Substituting (4.41) in (5.23) yields an identity between terminating balanced ${}_4F_3(1)$ series :

$$\begin{aligned}
&\sum_K f\,{}_4F_3\left({k_1+k_2+k_{12}-1,k_2+k_3+K-1,k_1+k_2-k_{12},k_2+k_3-K \atop 2k_2,k_{123}+k_1+k_2+k_3-1,k_1+k_2+k_3-k_{123}};1\right) \\
&\times {}_4F_3\left({k_1+K+k_{123}-1,K+k_4+k_{234}-1,k_1+K-k_{123},K+k_4-k_{234} \atop k_1+K+k_4-k,2K,k+k_1+K+k_4-1};1\right) \\
&\times {}_4F_3\left({k_2+k_3+K-1,k_2+k_3-K,k_3+k_4+k_{34}-1,k_3+k_4-k_{34} \atop k_2+k_3+k_4-k_{234},2k_3,k_{234}+k_2+k_3+k_4-1};1\right) = \\
&{}_4F_3\left({k_3+k_4+k_{34}-1,k_{12}+k_3+k_{123}-1,k_3+k_4-k_{34},k_{12}+k_3-k_{123} \atop k_{12}+k_3+k_4-k,2k_3,k+k_{12}+k_3+k_4-1};1\right) \\
&\times {}_4F_3\left({k_1+k_2+k_{12}-1,k_1+k_2-k_{12},k_2+k_{34}+k_{234}-1,k_2+k_{34}-k_{234} \atop 2k_2,k_1+k_2+k_{34}-k,k+k_1+k_2+k_{34}-1};1\right),
\end{aligned}$$

where f is a numerical factor that is easily calculated from (4.41). Renaming the following positive integer differences as

$$m=k_{12}-k_1-k_2,\ j=k_{123}-k_{12}-k_3 \text{ and } l=K-k_2-k_3,$$

and performing appropriate transformations (2.10) on the balanced ${}_4F_3(1)$'s, yields that (when $k_{123}-k_1\le k_{234}-k_4$) :

$$
\begin{aligned}
&\sum_{l=0}^{j+m} f' {}_4F_3\left(\genfrac{}{}{0pt}{}{-l,l+2k_2+2k_3-1,-m,m+2k_1+2k_2-1}{2k_2,2k_1+2k_2+2k_3+j+m-1,-j-m};1\right)\\
&\times {}_4F_3\left(\genfrac{}{}{0pt}{}{-m-j+l,m+j+l+2k_1+2k_2+2k_3-1,k_1+k+k_{234}-1,k_1+k-k_{234}}{2k_1,k_1+k_2+k_3+l+k_4+k-1,k_1+k_2+k_3+l+k-k_4};1\right)\\
&\times {}_4F_3\left(\genfrac{}{}{0pt}{}{-l,l+2k_2+2k_3-1,k_2+k_{234}+k_{34}-1,k_2+k_{234}-k_{34}}{2k_2,k_2+k_3+k_{234}-k_4,k_{234}+k_2+k_3+k_4-1};1\right)=\\
&{}_4F_3\left(\genfrac{}{}{0pt}{}{-j,j+2m+2k_1+2k_2+2k_3-1,m+k_1+k_2+k-k_{34},m+k_1+k_2+k+k_{34}-1}{2m+2k_1+2k_2,m+k_1+k_2+k_3+k_4+k-1,m+k_1+k_2+k_3+k-k_4};1\right)\\
&\times {}_4F_3\left(\genfrac{}{}{0pt}{}{-m,m+2k_1+2k_2-1,k_1+k+k_{234}-1,k_1+k-k_{234}}{2k_1,k_1+k_2+k-k_{34},k_1+k_2+k+k_{34}-1};1\right),
\end{aligned}
$$

where, once again, f' is a numerical factor that is easily calculated.

This last identity can be written in terms of Wilson and Racah polynomials by putting

$$
ix_1 = k - k_{234}, \qquad ix_2 = k_{234} - k_{34}, \qquad ix_3 = k_{34} - k_4
$$
$$
\text{and } it = k_4 + i(x_1 + x_2 + x_3) - 1/2.
$$

Note that all the renamings are invertible. Determination of the factor f' now yields the desired result. Since (5.24) is a rational identity in the parameters k_i, x_i and t, it is valid for all values of these parameters.

From Theorem 5.3 we can easily rederive a convolution identity for continuous Hahn polynomials, using the limit transition (5.21) :

Corollary 5.1. *The continuous Hahn polynomials satisfy the following convolution identity :*

$$
\begin{aligned}
&\sum_{l=0}^{m+j} \binom{j+m}{m} \frac{(2k_2)_m (2k_3)_j (2k_1+2k_2+2k_3+j+m-1)_l}{(2k_3)_l (2k_2+2k_3+l-1)_l (2k_2+2k_3+2l)_{j+m-l}} \qquad (5.25)\\
&\quad \times R_l(\lambda(m); 2k_2-1, 2k_3-1, -j-m-1, 2k_1+2k_2+j+m-1)\\
&\quad \times p_{m+j-l}(x_1; k_1, k_2+k_3+l-is, k_1, k_2+k_3+l+is)\\
&\quad \times p_l(x_2; k_2, k_3-i(s-x_1), k_2, k_3+i(s-x_1))\\
&= p_m(x_1; k_1, k_2-i(x_1+x_2), k_1, k_2+i(x_1+x_2))\\
&\quad \times p_j(x_1+x_2; k_1+k_2+m, k_3-is, k_1+k_2+m, k_3+is),
\end{aligned}
$$

where $j, m \in \mathbb{N}$, k_1, k_2, k_3, x_1, x_2, $x_3 \in \mathbb{R}$ *and* $s = x_1 + x_2 + x_3$.

This corollary (and the next one) has many interpretations, see [26]. In particular, using generalized eigenvectors of an operator $X_\varphi \in \mathfrak{su}(1,1)$ in $\mathfrak{su}(1,1)$ representations (rather than the eigenvectors of J_0 as used here), the continuous Hahn polynomials have an interpretation as Clebsch-Gordan coefficients for $\mathfrak{su}(1,1)$. Then Corollary 5.1 is the analog of (4.13), see also Exercise 4 of this section.

Finally, replacing x_i by sx_i in (5.25) and letting s tend to infinity yields Corollary 5.2 (see also [26, Corollary 3.15], or [17],[47]) :

Corollary 5.2. *The Jacobi polynomials satisfy the following convolution identity :*

$$
\begin{aligned}
&\sum_{l=0}^{m+j} \binom{j+m}{m} \frac{(2k_2)_m (2k_3)_j (2k_1+2k_2+2k_3+j+m-1)_l}{(2k_3)_l (2k_2+2k_3+l-1)_l (2k_2+2k_3+2l)_{j+m-l}} \\
&\quad \times R_l(\lambda(m); 2k_2-1, 2k_3-1, -j-m-1, 2k_1+2k_2+j+m-1) \\
&\quad \times P_{m+j-l}^{(2k_1-1, 2k_2+2k_3+2l-1)}(1-2x_1)(1-x_1)^l P_l^{(2k_2-1,2k_3-1)}\left(\frac{1-x_1-2x_2}{1-x_1}\right) \\
&= (x_1+x_2)^m P_m^{(2k_1-1,2k_2-1)}\left(\frac{x_2-x_1}{x_1+x_2}\right) \\
&\qquad\qquad \times P_j^{(2k_1+2k_2+2m-1,2k_3-1)}(1-2x_1-2x_2).
\end{aligned}
$$

This corollary can be written in a more symmetric (and unified) way, by defining [17], [47] :

$$
\begin{aligned}
S_m^{k_1,k_2}(x_1,x_2) =& (-1)^m \sqrt{\frac{m!}{(2k_1, 2k_2, 2k_1+2k_2+m-1)_m}} \\
& \times (x_1+x_2)^m P_m^{(2k_1-1,2k_2-1)}\left(\frac{x_2-x_1}{x_2+x_1}\right).
\end{aligned} \tag{5.26}
$$

Then corollary 5.2 reads ($x_1+x_2+x_3=1$) :

$$
\begin{aligned}
&S_{k_{12}-k_1-k_2}^{k_1,k_2}(x_1,x_2) S_{k_0-k_{12}-k_3}^{k_{12},k_3}(x_1+x_2,x_3) = \\
&\sum_{K=k_2+k_3}^{k_0-k_1} U_{k_3,k_0,K}^{k_1,k_2,k_{12}} S_{K-k_2-k_3}^{k_2,k_3}(x_2,x_3) S_{k_0-k_1-K}^{k_1,K}(x_1, x_2+x_3).
\end{aligned} \tag{5.27}
$$

Formula (5.27) is easily remembered by considering two ways in which three $\mathfrak{su}(1,1)$ representations can be coupled, as shown in Figure 6. Notice how the left side of (5.27) follows from the tree on the left side of this figure. With each non-leaf node (i.e. with each intermediate or final coupling) one associates an S-polynomial. The first (resp. second) variable of this S-polynomial is the sum of all the variables associated with the leaves in the left (resp. right) subtree of the considered node. The upper parameters are determined by the value of the representation labels of the left and right child (in that order). The (positive integer) lower parameter is the difference between the value of the coupled representation label and the consisting labels. The S-polynomials on the right side of (5.27) are formed in the same way but working with the tree on the right side of the figure. The recoupling coefficient appearing in (5.27) is that associated with a recoupling of three representations as shown in Figure 6; it can be seen as a manifestation of Rule 2, cfr. (5.5).

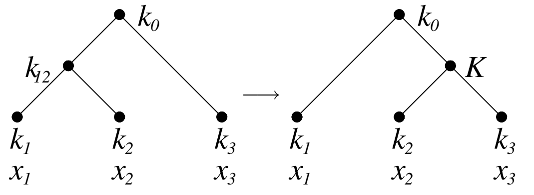


Fig. 6. Two possible ways of coupling three representations

The S-polynomials also have the following property :

$$S_m^{k_1,k_2}(x_1,x_2) = (-1)^m S_m^{k_2,k_1}(x_2,x_1). \tag{5.28}$$

This can be seen as a manifestation of Rule 1, (5.4).

The products of Jacobi polynomials in both the left and right side of (5.27) are orthogonal [26] on the simplex determined by $x_1, x_2 > 0$, $x_1 + x_2 < s$, for the weight function

$$w(x_1,x_2) = x_1^{2k_1-1} x_2^{2k_2-1} (s - x_1 - x_2)^{2k_3-1}.$$

So the two ways of taking tensor products of three $\mathfrak{su}(1,1)$ representations, yield two 2-variable Jacobi polynomials orthogonal with respect to $w(x_1,x_2)$, and related to one another by Racah coefficients of $\mathfrak{su}(1,1)$. This idea will be generalized to $(n+1)$-fold tensor products.

When considering orthogonal polynomials in n variables, one of the classical areas is the simplex T_s^n :

$$T_s^n = \{x \in \mathbb{R}^n \mid 0 < x_j \text{ and } |x| = x_1 + \cdots + x_n < s\}. \tag{5.29}$$

Herein, s denotes some positive constant, and in almost all cases s is taken to be equal to 1. The classical weight function in this case is :

$$x_1^{\kappa_1 - 1/2} \cdots x_n^{\kappa_n - 1/2} (s - |x|)^{\kappa_{n+1} - 1/2}, \tag{5.30}$$

where each $\kappa_i > -1/2$. In [13, Proposition 2.3.8] an explicit orthonormal basis is given associated with the weight function (5.30) on the simplex (5.29). Such a basis is not unique. In fact, with every binary coupling tree on $n+1$ leaves, a different basis can be constructed. In this section, an n-variable orthonormal polynomial will be constructed out of a product of n S-polynomials (5.26), and associated with a binary coupling scheme of $n+1$ representations of $\mathfrak{su}(1,1)$. It can be shown that this polynomial is orthogonal on the simplex T_s^n for the classical weight function.

Theorem 5.4. *With every coupling of* $(n+1)$ $\mathfrak{su}(1,1)$ *representations labelled by* $k_1, \ldots, k_{n+1}$, *i.e. with every binary coupling scheme with* n *internal nodes, we associate a set of polynomials*

$$R_l^{(k)}(x) \equiv R_{(l_1,\ldots,l_n)}^{(k_1,\ldots,k_{n+1})}(x_1,\ldots,x_n)$$

in n variables orthogonal on the simplex T_s^n for the weight function

$$w^{(k)}(x) = x_1^{2k_1-1}\cdots x_n^{2k_n-1}(s-|x|)^{2k_{n+1}-1}, \tag{5.31}$$

where each $k_i > 0$. Explicitly the orthogonality reads :

$$\int_{T_s^n} R_l^{(k)}(x) R_{l'}^{(k)}(x) w^{(k)}(x)\, dx = \delta_{l,l'} \frac{s^{2|k|+2|l|-1}}{\Gamma(2|k|+2|l|)} \prod_{i=1}^{n+1} \Gamma(2k_i). \tag{5.32}$$

In principal the notation of the polynomial should contain a reference to the binary coupling scheme it corresponds to. For the moment, we can assume that the binary coupling scheme is fixed, and we do not mention it in the notation of $R_l^{(k)}(x)$. When we want to emphasize the dependence of $R_l^{(k)}(x)$ on the given binary coupling scheme T, we shall write $R_{l,T}^{(k)}(x)$. The meaning of the subscript l is related to the labelling of the internal nodes, and will soon become apparent.

The association of a polynomial with a binary coupling scheme is an extension of the method described after Eq. (5.27). For a given binary coupling scheme, the polynomial $R_l^{(k)}(x)$ consists of a product of S-polynomials, each of these associated with a non-leaf node of the tree.

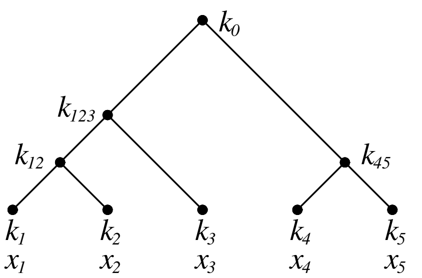


Fig. 7. Example binary coupling tree

Let us first describe an example. With the binary coupling scheme shown in Figure 7 we associate the following polynomial :

$$
\begin{aligned}
&R(x_1,x_2,x_3,x_4,x_5)\\
&= S^{k_1,k_2}_{k_{12}-k_1-k_2}(x_1,x_2)S^{k_{12},k_3}_{k_{123}-k_{12}-k_3}(x_1+x_2,x_3)S^{k_4,k_5}_{k_{45}-k_4-k_5}(x_4,x_5)\\
&\quad\times S^{k_{123},k_{45}}_{k_0-k_{123}-k_{45}}(x_1+x_2+x_3,x_4+x_5)\\
&= C(x_1+x_2)^{k_{12}-k_1-k_2}P^{(2k_1-1,2k_2-1)}_{k_{12}-k_1-k_2}\left(\frac{x_2-x_1}{x_2+x_1}\right)\\
&\quad\times(x_1+x_2+x_3)^{k_{123}-k_{12}-k_3}P^{(2k_{12}-1,2k_3-1)}_{k_{123}-k_{12}-k_3}\left(\frac{x_3-x_1-x_2}{x_3+x_1+x_2}\right)\\
&\quad\times(x_4+x_5)^{k_{45}-k_4-k_5}P^{(2k_4-1,2k_5-1)}_{k_{45}-k_4-k_5}\left(\frac{x_5-x_4}{x_5+x_4}\right)\\
&\quad\times(x_1+x_2+x_3+x_4+x_5)^{k_0-k_{123}-k_{45}}\\
&\qquad\qquad\times P^{(2k_{123}-1,2k_{45}-1)}_{k_0-k_{123}-k_{45}}\left(\frac{x_4+x_5-x_1-x_2-x_3}{x_4+x_5+x_1+x_2+x_3}\right),
\end{aligned}
$$

herein, C is some numerical factor, that can be determined from (5.26).

In general, the S-polynomial, associated to a non-leaf node of the tree, has : as (upper) parameters the representation labels of left and right child of the node; as degree (the sub-index) the difference between the representation label of the node and those of the children (this is a nonnegative integer); as left (resp. right) argument the sum of all the variables associated with the leaves in the left (resp. right) subtree of the considered node.

Such a polynomial $R(x_1,\ldots,x_{n+1})$, defined as a product of S-polynomials in this way, is homogeneous in the variables $x_1,\ldots,x_{n+1}$. So we can choose the constraint :

$$x_1+x_2+\cdots+x_{n+1}=s, \tag{5.33}$$

where s is some arbitrary, positive constant. Note that this constraint is compatible with the definition of the weight function (5.31). The resulting polynomial will be denoted by $R^{(k)}_l(x)$. The subscript l in $R^{(k)}_l(x)$ stands for the sequence of degrees of the S-polynomials, in a chosen order.

The proof of Theorem 5.4 can be given by an explicit change of variables [29]. It can be extended in the following way :

Theorem 5.5. *Consider a binary coupling scheme, T_1, with fixed values k_j and l_i. Consider another binary coupling scheme T_2 with the same fixed values k_j but varying values l'_i, such that $|l|=|l'|$. Then the polynomials $R^{(k)}_{l,T_1}(x)$ can be written as a linear combination of polynomials $R^{(k)}_{l',T_2}(x)$:*

$$R^{(k)}_{l,T_1}(x)=\sum_{|l'|=|l|}C_{l'}R^{(k)}_{l',T_2}(x). \tag{5.34}$$

The connection coefficient $C_{l'}$ is equal to the $3nj$-coefficient $\langle T_1,T_2\rangle$ (which is zero anyway if $|l|\neq|l'|$).

So for two n-variable Jacobi polynomials corresponding to the same binary coupling scheme, their inner product is given by (5.32). For two n-variable Jacobi polynomials with different binary coupling schemes, the inner product is essentially given by a $3nj$-coefficient :

$$\int_{T_s^n} R_{l,T_1}^{(k)}(x) R_{l',T_2}^{(k)}(x) w^{(k)}(x) dx = \langle T_1(l), T_2(l') \rangle \frac{s^{2|k|+2|l|-1}}{\Gamma(2|k|+2|l|)} \prod_{i=1}^{n+1} \Gamma(2k_i), \tag{5.35}$$

where $w^{(k)}(x)$ is the classical weight function (5.31).

5.6 Notes and Exercises

The $3nj$-coefficients have been considered mainly in the physics literature, but only for particular n-values. The $9j$-coefficients (more precisely, a particular $9j$-coefficient, if one follows the terminology of these notes) appeared already in Wigner's paper [52], and were studied by Jahn and Hope [21]. They also introduced $12j$-coefficients. The idea of binary coupling in general appears in the book of Biedenharn and Louck [7, Topic 12, volume 9], for the $\mathfrak{su}(2)$ case. Here we have shown that this technique can be applied to the $\mathfrak{su}(1,1)$ case as well. In this context, it should be mentioned that a powerful graphical technique was developed by the Jucys school, in order to reduce and compute $3nj$-coefficients [23]. Their technique, however, only applies to the $\mathfrak{su}(2)$ case.

The Biedenharn-Elliott idendity [6][10] was originally just a curious identity, also useful to derive recurrence relations between Racah coefficients. Its essential role in mathematical applications became clear only much later, see e.g. [12], especially with the developments of quantum groups.

The technique of binary coupling trees was used in [11], and described in [28]. Applications and implementations are given in [14]. A systematic study of binary coupling trees, the graph G_n, and optimal expressions for $3nj$-coefficients is found in [15], [16].

The convolution identities of section 5.5 find their background in [26]. Here, they are deduced differently, by formal substitutions in the Biedenharn-Elliott identity. The multivariable orthogonal polynomials of this section were analysed in [29]; some special cases appeared already in the literature, see references in [29]. The relation between tensor products of positive discrete series representations of $\mathfrak{su}(1,1)$ and multivariable orthogonal polynomials is also discussed in detail in the work of Rosengren [40].

Exercises

1. Use the technique of binary coupling to show that

$$U_{j_3,j,j_{23}}^{j_1,j_2,j_{12}} = \sum_l (-1)^{j_1+j_2+j_3+j-j_{12}-j_{23}-l} U_{j_3,j,l}^{j_2,j_1,j_{12}} U_{j_2,j,j_{23}}^{j_1,j_3,l}.$$

Rewrite this in the form

$$\sum_x (-1)^{p+q+x}(2x+1) \begin{Bmatrix} a & b & x \\ c & d & p \end{Bmatrix} \begin{Bmatrix} a & b & x \\ d & c & q \end{Bmatrix} = \begin{Bmatrix} a & c & q \\ b & d & p \end{Bmatrix}. \tag{5.36}$$

2. Assume that (a, b, c) forms a triad. Show that

$$\sum_x (2x+1) \begin{Bmatrix} a & b & x \\ a & b & c \end{Bmatrix} = (-1)^{2c}, \tag{5.37}$$

   where the sum is over all (integer or half-integer) x which obey all triangular conditions.
3. Let j_1, j_2, j_3 and j_4 be four $\mathfrak{su}(2)$ representation labels. In the context of angular momentum theory one defines the $9j$-coefficient as :

$$\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j_4 & j_{34} \\ j_{13} & j_{24} & j \end{Bmatrix} = \frac{\langle e_m^{((j_1 j_3)j_{13}(j_2 j_4)j_{24})j}, e_m^{((j_1 j_2)j_{12}(j_3 j_4)j_{34})j}\rangle}{\sqrt{(2j_{12}+1)(2j_{34}+1)(2j_{13}+1)(2j_{24}+1)}}.$$

   Apply the technique of binary recoupling trees to compute an expansion of this coefficient in terms of Racah coefficients. Deduce that

$$\sum_x (-1)^{2x}(2x+1) \begin{Bmatrix} a & b & x \\ c & d & p \end{Bmatrix} \begin{Bmatrix} c & d & x \\ e & f & q \end{Bmatrix} \begin{Bmatrix} e & f & x \\ a & b & r \end{Bmatrix} = \begin{Bmatrix} a & f & r \\ d & q & e \\ p & c & b \end{Bmatrix}. \tag{5.38}$$

   Observe the similarities between (5.37) and (4.44), (4.43) and (5.36), (5.38) and (5.13).
4. Verify that by choosing the parameters in Corollary 5.1 so that the continuous Hahn polynomials become the (discrete) Hahn polynomials, one obtains a convolution identity for Hahn polynomials. Furthermore, check that the identity thus obtained is equivalent with (4.13).

Acknowledgements

I would like to thank Stijn Lievens for a careful reading of the manuscript.

References

1. G.E. Andrews, R. Askey and R. Roy, *Special functions*. Cambridge University Press, Cambridge, 1999.
2. R. Askey and J.A. Wilson, A set of orthogonal polynomials that generalize the Racah coefficients of 6-j symbols, *SIAM J. Math. Anal.* **10** (1979), 1008–1016.
3. R. Askey and J.A. Wilson, Some basic hypergeometric orthogonal polynomials that generalize Jacobi polynomials, *Mem. Amer. Math. Soc.* **54** (1985), no. 319.
4. W.N. Bailey, *Generalized hypergeometric series*. Cambridge University Press, Cambridge, 1935.

5. V. Bargmann, On the representations of the rotation group, *Rev. Mod. Phys.* **34** (1962), 829–845.
6. L.C. Biedenharn, An identity satisfied by the Racah coefficients, *J. Math. Phys.* **31** (1953), 287–293.
7. L.C. Biedenharn and J.D. Louck, *Angular momentum in quantum physics.* Encyclopedia of mathematics and its applications, vol. 8. Addison-Wesley, Reading, 1981.
   The Racah-Wigner algebra in quantum theory. Encyclopedia of mathematics and its applications, vol. 9. Addison-Wesley, Reading, 1981.
8. L.C. Biedenharn and H. Van Dam, *Quantum theory of angular momentum.* Academic Press, New York, 1965.
9. N. Bourbaki, *Groupes et algèbres de Lie.* Chap. 1, 4–6, 7–8. Hermann, Paris, 1960, 1968, 1975.
10. J.P. Elliott, Theoretical studies in nuclear structure V : The matrix elements of non-central forces with an application to the $2p$-shell, *Proc. Roy. Soc.* **A218** (1953), 370.
11. P.G. Burke, A program to calculate a general recoupling coefficient, *Comput. Phys. Commun.* **1** (1970), 241–250.
12. J.S. Carter, D.E. Flath and M. Saito, *The classical and quantum* $6j$ *-symbols.* Princeton University Press, Princeton, 1995.
13. C.F. Dunkl and Y. Xu, *Orthogonal polynomials of several variables.* Encyclopedia of mathematics and its applications, vol. 81. Cambridge Univ. Press, 2001.
14. V. Fack, S.N. Pitre and J. Van der Jeugt, New efficient programs to calculate general recoupling coefficients, I, *Comput. Phys. Commun.* **83** (1994), 275–292.
15. V. Fack, S. Lievens and J. Van der Jeugt, On rotation distance between binary coupling trees and applications for $3nj$-coefficients, *Comput. Phys. Commun.* **119** (1999), 99–114.
16. V. Fack, S. Lievens and J. Van der Jeugt, On the diameter of the rotation graph of binary coupling trees, *Discrete Mathematics* **245** (2002), 1–18.
17. Y.I. Granovskiĭ and A.S. Zhedanov, New construction of $3nj$-symbols, *J. Phys. A : Math. Gen.* **26** (1993), 4339–4344.
18. W.J. Holman and L.C. Biedenharn, A general study of the Wigner coefficients of $SU(1,1)$, *Ann. Physics* **47** (1968), 205–231.
19. J.E. Humphreys, *Introduction to Lie algebras and representation theory.* Springer Verlag, Berlin, 1972.
20. N.A. Jacobson, *Lie algebras.* Wiley, New York, 1962.
21. H.A. Jahn and J. Hope, Symmetry properties of the Wigner $9j$ symbol, *Phys. Rev.* **93** (1954), 318–321.
22. J.C. Jantzen, *Moduln mit einem höchsten Gewicht.* Lecture Notes in Mathematics, 750. Springer, Berlin, 1979.
23. A.P. Jucys, I.B. Levinson and V.V. Vanagas, *Mathematical apparatus of the theory of angular momentum.* Russian edition, 1960. English translation : Israel program for scientific translations, Jerusalem, 1962.
24. S. Karlin and J.L. McGregor, The Hahn polynomials, formulas and an application, *Scripta Math.* **26** (1961), 33–46.
25. R. Koekoek and R.F. Swarttouw, *The Askey-scheme of hypergeometric orthogonal polynomials and its* q *-analogue.* Report 98-17, Technical University Delft (1998); available at `http://aw.twi.tudelft.nl/~koekoek/askey.html`.

26. H.T. Koelink and J. Van der Jeugt, Convolutions for orthogonal polynomials from Lie and quantum algebra representations, *SIAM J. Math. Anal.* **29** (1998), 794–822.
27. T.H. Koornwinder, Clebsch-Gordan coefficients for SU(2) and Hahn polynomials, *Nieuw Archief voor Wiskunde* (3), **XXIX** (1981), 140–155.
28. G.I. Kuznetsov and Ya. Smorodinskiĭ, On the theory of $3nj$ coefficients, *Soviet J. Nuclear Phys.* **21** (1975), 585–588.
29. S. Lievens and J. Van der Jeugt, $3nj$-coefficients of $su(1,1)$ as connection coefficients between orthogonal polynomials in n variables, *J. Math. Phys.* **43** (2002), 3824–3849.
30. S.D. Majumdar, The Clebsch-Gordan coefficients, *Prog. Theor. Phys.* **20** (1958), 798–803.
31. J. von Neumann and E. Wigner, Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons, III, *Z. Physik* **51** (1928), 844–858.
32. W. Pauli, Zur Quantenmechanik des magnetischen Elektrons, *Z. Physik* **43** (1927), 601–623.
33. G. Racah, Theory of complex spectra. I, II, III and IV. *Phys. Rev.* **61** (1942), 186–197; **62** (1942), 438–462; **63** (1943), 367–382; **76** (1949), 1352–1365.
34. W. Rasmussen, Identity of the $SU(1,1)$ and $SU(2)$ Clebsch-Gordan coefficients coupling unitary discrete representations, *J. Phys. A* **8** (1975), 1038–1047.
35. K. Srinivasa Rao, T.S. Santhanam and K. Venkatesh, On the symmetries of the Racah coefficient, *J. Math. Phys.* **16** (1975), 1528–1530.
36. K. Srinivasa Rao and V. Rajeswari, *Quantum theory of angular momentum. Selected topics.* Narosa/Springer , New Delhi, 1993.
37. T. Regge, Symmetry properties of Clebsch-Gordan's coefficients, *Nuovo Cimento* **10** (1958), 544–545.
38. T. Regge, Symmetry properties of Racah's coefficients, *Nuovo Cimento* **11** (1959), 116–117.
39. J. Repka, Tensor products of unitary representations of $SL_2(R)$, *Amer. J. Math.* **100** (1978), 747–774.
40. H. Rosengren, Multilinear Hankel forms of higher order and orthogonal polynomials, *Math. Scand.* **82** (1998), 53–88.
    Multivariable orthogonal polynomials and coupling coefficients for discrete series representations, *SIAM J. Math. Anal.* **30** (1999), 232–272.
41. J. Schwinger, On angular momentum, unpublished (1952); see [8]
42. J.-P. Serre, *Lie algebras and Lie groups.* W.A. Benjamin, New York, 1966.
43. L.J. Slater, *Generalized hypergeometric functions.* Cambridge Univ. Press, Cambridge, 1966.
44. Ya. Smorodinskiĭ and L.A. Shelepin, Clebsch-Gordan coefficients, viewed from different sides, *Sov. Phys. Uspekhi* **15** (1972), 1–24.
45. Ya. Smorodinskiĭ and S.K. Suslov, The Clebsch-Gordan coefficients of the group $SU(2)$ and Hahn polynomials, *Soviet J. Nuclear Phys.* **35** (1982), 192–201.
46. H. Ui, Clebsch-Gordan formulas of the $SU(1,1)$ group, *Progr. Theoret. Phys.* **44** (1970), 689–702.
47. J. Van der Jeugt, Coupling coefficients for Lie algebra representations and addition formulas for special functions, *J. Math. Phys.* **38** (1997), 2728–2740.
48. J. Van der Jeugt and K. Srinivasa Rao, Invariance groups of transformations of basic hypergeometric series, *J. Math. Phys.* **40** (1999), 6692–6700.

49. V.S. Varadarajan, *Lie groups, Lie algebras, and their representations*. Prentice-Hall, New York, 1974.
50. D.A. Varshalovich, A.N. Moskalev and V.K. Khersonskii, *Quantum theory of angular momentum*. Nauka, Leningrad, 1975 : in Russian. English Edition : World Scientific, Singapore, 1988.
51. N.Ja. Vilenkin and A.U. Klimyk, *Representation of Lie groups and special functions*, vols. 1, 2 and 3. Kluwer Academic Press, Dordrecht, 1991, 1992.
52. E.P. Wigner, On the matrices which reduce the Kronecker products of representations of S.R. groups, unpublished (1940); see [8].
53. E.P. Wigner, *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspekren.* Vieweg, Braunschweig, 1931.
    Group theory and its applications to the quantum mechanics of atomic spectra. Academic Press, New York, 1959.
54. J.A. Wilson, Some hypergeometric orthogonal polynomials, *SIAM J. Math. Anal.* **11** (1980), 690–701.

Dunkl Operators: Theory and Applications

Margit Rösler

Mathematisches Institut, Universität Göttingen, Bunsenstr. 3-5, D-37073 Göttingen, Germany, roesler@uni-math.gwdg.de

Summary. These lecture notes are intended as an introduction to the theory of rational Dunkl operators and the associated special functions, with an emphasis on positivity and asymptotics. We start with an outline of the general concepts: Dunkl operators, the intertwining operator, the Dunkl kernel and the Dunkl transform. We point out the connection with integrable particle systems of Calogero-Moser-Sutherland type, and discuss some systems of orthogonal polynomials associated with them. A major part is devoted to positivity results for the intertwining operator and the Dunkl kernel, the Dunkl-type heat semigroup, and related probabilistic aspects. The notes conclude with recent results on the asymptotics of the Dunkl kernel.

1 Introduction ... 94

2 Dunkl operators and the Dunkl transform ... 95

2.1 Root systems and reflection groups ... 95
2.2 Dunkl operators ... 97
2.3 A formula of Macdonald and its analog in Dunkl theory ... 100
2.4 Dunkl's intertwining operator ... 103
2.5 The Dunkl kernel ... 105
2.6 The Dunkl transform ... 109

3 CMS models and generalized Hermite polynomials ... 110

3.1 Quantum Calogero-Moser-Sutherland models ... 110
3.2 Spectral analysis of abstract CMS Hamiltonians ... 112
3.3 Generalized Hermite polynomials ... 114

4 Positivity results ... 117

4.1 Positivity of Dunkl's intertwining operator ... 117
4.2 Heat kernels and heat semigroups ... 120

5 Asymptotic analysis for the Dunkl kernel ... 127

6 Notation ... 132

References ... 132

E. Koelink and W. Van Assche (Eds.): LNM 1817, pp. 93–135, 2003.

1 Introduction

While the theory of special functions in one variable has a long and rich history, the growing interest in special functions of several variables is comparatively recent. During the last years, there has in particular been a rapid development in the area of special functions with reflection symmetries and the harmonic analysis related with root systems. The motivation for this subject comes to some extent from the theory of Riemannian symmetric spaces, whose spherical functions can be written as multi-variable special functions depending on certain discrete sets of parameters. A key tool in the study of special functions with reflection symmetries are Dunkl operators. Generally speaking, these are commuting differential-difference operators, associated to a finite reflection group on a Euclidean space. The first class of such operators, now often called "rational" Dunkl operators, were introduced by C.F. Dunkl in the late 80ies. In a series of papers ([11]-[15]), he built up the framework for a theory of special functions and integral transforms in several variables related with reflection groups. Since then, various other classes of Dunkl operators have become important, in the first place the trigonometric Dunkl operators of Heckman, Opdam and the Cherednik operators. These will not be discussed in our notes; for an overview, we refer to [27]. An important motivation to study Dunkl operators originates in their relevance for the analysis of quantum many body systems of Calogero-Moser-Sutherland type. These describe algebraically integrable systems in one dimension and have gained considerable interest in mathematical physics, especially in conformal field theory. A good bibliography is contained in [10].

The aim of these lecture notes is an introduction to rational Dunkl theory, with an emphasis on the author's results in this area. Rational Dunkl operators bear a rich analytic structure which is not only due to their commutativity, but also to the existence of an intertwining operator between Dunkl operators and usual partial derivatives. We shall first give an overview of the general concepts, including an account on the relevance of Dunkl operators in the study of Calogero-Moser-Sutherland models. We also discuss some of the special functions related with them. A major topic will be positivity results; these concern the intertwining operator as well as the kernel of the Dunkl transform, and lead to a variety of positive semigroups in the Dunkl setting with possible probabilistic interpretations. We make this explicit at hand of the most important example: the Dunkl-type heat semigroup, which is generated by the analog of the Laplacian in the Dunkl setting. The last section presents recent results on the asymptotics of the Dunkl kernel and the short-time behavior of heat kernels associated with root systems.

2 Dunkl operators and the Dunkl transform

The aim of this section is to provide an introduction to the theory of rational Dunkl operators, which we shall call Dunkl operators for short, and to the Dunkl transform. General references are [11]-[15], [18], [30] and [44]; for a background on reflection groups and root systems the reader is referred to [29] and [23]. We do not intend to give a complete survey, but rather focus on those aspects which will be important in the context of this lecture series.

2.1 Root systems and reflection groups

The basic ingredient in the theory of Dunkl operators are root systems and finite reflection groups, acting on some Euclidean space $(E, \langle \,.\,,.\, \rangle)$ of finite dimension N. We shall always assume that $E = \mathbb{R}^N$ with the standard Euclidean scalar product $\langle x, y\rangle = \sum_{j=1}^N x_j y_j$. For $\alpha \in \mathbb{R}^N \setminus \{0\}$, we denote by σ_α the reflection in the hyperplane $\langle \alpha\rangle^\perp$ orthogonal to α, i.e.

$$\sigma_\alpha(x) = x - 2\,\frac{\langle \alpha, x\rangle}{|\alpha|^2}\,\alpha\,,$$

where $|x| := \sqrt{\langle x, x\rangle}$. Each reflection σ_α is contained in the orthogonal group $O(N,\mathbb{R})$. We start with the basic definitions:

Definition 2.1. *Let $R \subset \mathbb{R}^N \setminus \{0\}$ be a finite set. Then R is called a root system, if*

(1) *$R \cap \mathbb{R}\alpha = \{\pm\alpha\}$ for all $\alpha \in R$;*
(2) *$\sigma_\alpha(R) = R$ for all $\alpha \in R$.*

The subgroup $G = G(R) \subseteq O(N,\mathbb{R})$ which is generated by the reflections $\{\sigma_\alpha,\, \alpha \in R\}$ is called the reflection group (or Coxeter-group) associated with R. The dimension of $span_{\mathbb{R}} R$ is called the rank of R.

Property (1) is called reducedness. It is often not required in Lie-theoretic contexts, where instead the root systems under consideration are assumed to be crystallographic. This means that

$$\frac{2\langle \alpha, \beta\rangle}{\langle \beta, \beta\rangle} \in \mathbb{Z} \quad \text{for all } \alpha, \beta \in R.$$

If R is crystallographic and has full rank, then $span_{\mathbb{Z}} R$ forms a lattice in $\mathbb{R}^N$ (called the *root-lattice*) which is stabilized by the action of the associated reflection group.

Lemma 2.1. (1) *For any root system R in $\mathbb{R}^N$, the reflection group $G = G(R)$ is finite.*
(2) *The set of reflections contained in $G(R)$ is exactly $\{\sigma_\alpha\,,\, \alpha \in R\}$.*

Proof. As R is left invariant by G, we have a natural homomorphism $\varphi : G \to S(R)$ of G into the symmetric group of R, given by $\varphi(g)(\alpha) := g\alpha \in R$. This homomorphism is injective: indeed, each reflection s_α, and therefore also each element $g \in G$ fixes pointwise the orthogonal complement of the subspace spanned by R. If also $g(\alpha) = \alpha$ for all $\alpha \in R$, then g must be the identity. This implies assertion (1) because the order of $S(R)$ is finite. Property (2) is more involved. An elegant proof can be found in Section 4.2 of [19].

Exercise 1. If $g \in O(N, \mathbb{R})$ and $\alpha \in \mathbb{R}^N \setminus \{0\}$, then $g\,\sigma_\alpha g^{-1} = \sigma_{g\alpha}$.

Together with part (2) of the previous lemma, this shows that there is a bijective correspondence between the conjugacy classes of reflections in G and the orbits in R under the natural action of G. We shall need some more concepts: Each root system can be written as a disjoint union $R = R_+ \cup (-R_+)$, where R_+ and $-R_+$ are separated by a hyperplane through the origin. Such a set R_+ is called a *positive subsystem.* Of course, its choice is not unique. The set of hyperplanes $\{\langle\alpha\rangle^\perp,\ \alpha \in R\}$ divides $\mathbb{R}^N$ into connected open components, called the *Weyl chambers* of R. It can be shown that the topological closure $\overline{C}$ of any chamber C is a fundamental domain for G, i.e. $\overline{C}$ is naturally homeomorphic with the space $(\mathbb{R}^N)^G$ of all G-orbits in $\mathbb{R}^N$, endowed with the quotient topology. G permutes the reflecting hyperplanes as well as the chambers.

Exercise 2. *Dihedral groups.* In the Euclidean plane $\mathbb{R}^2$, let $d \in O(2, \mathbb{R})$ denote the rotation around $2\pi/n$ with $n \geq 3$ and s the reflection at the y-axis. Show that the group $\mathcal{D}_n$ generated by d and s consists of all orthogonal transformations which preserve a regular n-sided polygon centered at the origin. (Hint: $dsd = s$.) Show that D_n is a finite reflection group and determine its root system. Can the crystallographic condition always be satisfied?

Examples 2.1. **(1) Type** A_{N-1}. Let S_N denote the symmetric group in N elements. It acts faithfully on $\mathbb{R}^N$ by permuting the standard basis vectors $e_1, \ldots, e_N$. Each transposition (ij) acts as a reflection σ_{ij} sending $e_i - e_j$ to its negative. Since S_N is generated by transpositions, it is a finite reflection group. A root system of S_N is given by

$$R = \{\pm(e_i - e_j),\ 1 \leq i < j \leq N\}.$$

Its span is $(e_1 + \ldots + e_N)^\perp$, and thus the rank is $N - 1$.

(2) Type B_N. Here G is the reflection group in $\mathbb{R}^N$ generated by the transpositions σ_{ij} as above, as well as the sign changes $\sigma_i : e_i \mapsto -e_i$, $i = 1, \ldots, N$. The group of sign changes is isomorphic to $\mathbb{Z}_2^N$, intersects S_N trivially and is normalized by S_N, so G is isomorphic with the semidirect product $S_N \ltimes \mathbb{Z}_2^N$. The corresponding root system has rank N; it is given by

$$R = \{\pm e_i,\ 1 \leq i \leq N,\ \pm(e_i \pm e_j), 1 \leq i < j \leq N\}.$$

A root system R is called *irreducible*, if it cannot be written as the orthogonal disjoint union $R = R_1 \cup R_2$ of two root systems R_1, R_2. Any root system can be uniquely written as an orthogonal disjoint union of irreducible root systems. There exists a classification of all irreducible root systems in terms of Coxeter graphs. The crystallographic ones are made up by 4 infinite families A_n, B_n (those discussed above), C_n, D_n, as well as 5 exceptional root systems. For details, we refer to [29].

2.2 Dunkl operators

Let R be a fixed root system in $\mathbb{R}^N$ and G the associated reflection group. From now on we assume that R is normalized in the sense that $\langle \alpha, \alpha \rangle = 2$ for all $\alpha \in R$; this simplifies formulas, but is no loss of generality for our purposes. The Dunkl operators attached with R can be considered as perturbations of the usual partial derivatives by reflection parts. These reflection parts are coupled by parameters, which are given in terms of a multiplicity function:

Definition 2.2. *A function $k : R \to \mathbb{C}$ on the root system R is called a multiplicity function on R, if it is invariant under the natural action of G on R. The $\mathbb{C}$-vector space of multiplicity functions on R is denoted by K.*

Notice that the dimension of K is equal to the number of G-orbits in R. We write $k \geq 0$ if $k(\alpha) \geq 0$ for all $\alpha \in R$.

Definition 2.3. *Let $k \in K$. Then for $\xi \in \mathbb{R}^N$, the Dunkl operator $T_\xi := T_\xi(k)$ is defined (for $f \in C^1(\mathbb{R}^N)$) by*

$$T_\xi f(x) := \partial_\xi f(x) + \sum_{\alpha \in R_+} k(\alpha) \langle \alpha, \xi \rangle \frac{f(x) - f(\sigma_\alpha x)}{\langle \alpha, x \rangle}.$$

Here ∂_ξ denotes the directional derivative corresponding to ξ, and R_+ is a fixed positive subsystem of R. For the i-th standard basis vector $\xi = e_i \in \mathbb{R}^N$ we use the abbreviation $T_i = T_{e_i}$.

The above definition does not depend on the special choice of R_+, thanks to the G-invariance of k. In case $k = 0$, the T_ξ reduce to the corresponding directional derivatives. The operators T_ξ were introduced and first studied for $k \geq 0$ by C.F. Dunkl ([11]-[15]). They enjoy regularity properties similar to usual partial derivatives on various spaces of smooth functions on $\mathbb{R}^N$. We shall use the following notations:

Notation 2.1. 1. $\mathbb{Z}_+ := \{0, 1, 2, \ldots\}$.

2. $\Pi := \mathbb{C}[\mathbb{R}^N]$ is the $\mathbb{C}$-algebra of polynomial functions on $\mathbb{R}^N$. It has a natural grading
$$\Pi = \bigoplus_{n \geq 0} \mathcal{P}_n ,$$
where $\mathcal{P}_n$ is the subspace of homogeneous polynomials of (total) degree n.

3. $\mathscr{S}(\mathbb{R}^N)$ denotes the Schwartz space of rapidly decreasing functions on $\mathbb{R}^N$,

$$\mathscr{S}(\mathbb{R}^N) := \{f \in C^\infty(\mathbb{R}^N) : \|x^\beta \partial^\alpha f\|_{\infty,\mathbb{R}^N} < \infty \quad \text{for all } \alpha, \beta \in \mathbb{Z}_+^N\}.$$

It is a Fréchet space with the usual locally convex topology.

The Dunkl operators T_ξ have the following regularity properties:

Lemma 2.2. (1) *If $f \in C^m(\mathbb{R}^N)$ with $m \geq 1$, then $T_\xi f \in C^{m-1}(\mathbb{R}^N)$.*
(2) *T_ξ leaves $C_c^\infty(\mathbb{R}^N)$ and $\mathscr{S}(\mathbb{R}^N)$ invariant.*
(3) *T_ξ is homogeneous of degree -1 on Π, that is, $T_\xi\, p \in \mathcal{P}_{n-1}$ for $p \in \mathcal{P}_n$.*

Proof. All statements follow from the representation

$$\frac{f(x) - f(\sigma_\alpha x)}{\langle \alpha, x\rangle} = \int_0^1 \partial_\alpha f(x - t\langle \alpha, x\rangle\alpha)dt \quad \text{for } f \in C^1(\mathbb{R}^N),\ \alpha \in R$$

(recall our normalization $\langle \alpha, \alpha\rangle = 2$). (1) and (3) are immediate; the proof of (2) (for $\mathscr{S}(\mathbb{R}^N)$) is also straightforward but more technical; it can be found in [30].

Due to the G-invariance of k, the Dunkl operators T_ξ are G-equivariant: In fact, consider the natural action of $O(N,\mathbb{R})$ on functions $f : \mathbb{R}^N \to \mathbb{C}$, given by

$$h \cdot f(x) := f(h^{-1}x), \quad h \in O(N,\mathbb{R}).$$

Then an easy calculation shows:

Exercise 3. $g \circ T_\xi \circ g^{-1} = T_{g\xi}$ for all $g \in G$.

Moreover, there holds a product rule:

Exercise 4. If f is G-invariant then $T_\xi f = \partial_\xi f$. If $f, g \in C^1(\mathbb{R}^N)$ and at least one of them is G-invariant, then

$$T_\xi(fg) = T_\xi(f) \cdot g + f \cdot T_\xi(g). \tag{2.1}$$

The most striking property of the Dunkl operators, which is the foundation for rich analytic structures related with them, is the following

Theorem 2.1. *For fixed k, the associated $T_\xi = T_\xi(k)$, $\xi \in \mathbb{R}^N$ commute.*

This result was obtained in [12] by a clever direct argumentation. An alternative proof, relying on Koszul complex ideas, is given in [18]. As a consequence of Theorem 2.1 there exists an algebra homomorphism $\Phi_k : \Pi \to \mathrm{End}_\mathbb{C}(\Pi)$ which is defined by

$$\Phi_k : x_i \mapsto T_i, \ 1 \mapsto id.$$

For $p \in \Pi$ we write

$$p(T) := \Phi_k(p).$$

The classical case $k = 0$ will be distinguished by the notation $\Phi_0(p) =: p(\partial)$. Of particular importance is the k-*Laplacian*, which is defined by

$$\Delta_k := p(T) \quad \text{with} \ \ p(x) = |x|^2.$$

Theorem 2.2.

$$\Delta_k \ = \Delta + 2 \sum_{\alpha \in R_+} k(\alpha)\delta_\alpha \qquad \textit{with} \qquad \delta_\alpha f(x) \ = \ \frac{\langle \nabla f(x), \alpha \rangle}{\langle \alpha, x \rangle} - \frac{f(x) - f(\sigma_\alpha x)}{\langle \alpha, x\rangle^2}; \tag{2.2}$$

here Δ and ∇ denote the usual Laplacian and gradient respectively.

This representation is obtained by a direct calculation (recall again our convention $\langle \alpha, \alpha \rangle = 2$ for all $\alpha \in R$) by use of the following Lemma:

Lemma 2.3. [12] *For $\alpha \in R$, define*

$$\rho_\alpha f(x) := \frac{f(x) - f(\sigma_\alpha x)}{\langle \alpha, x \rangle} \quad (f \in C^1(\mathbb{R}^N)).$$

Then

$$\sum_{\alpha, \beta \in R_+} k(\alpha)k(\beta)\langle \alpha, \beta \rangle \rho_\alpha \rho_\beta \ = \ 0.$$

It is not difficult to check that

$$\Delta_k = \sum_{i=1}^N T_{\xi_i}^2$$

for any orthonormal basis $\{\xi_1, \dots, \xi_N\}$ of $\mathbb{R}^N$, see [12] for the proof. Together with the G-equivariance of the Dunkl operators, this immediately implies that Δ_k is G-invariant, i.e.

$$g \circ \Delta_k = \Delta_k \circ g \quad (g \in G).$$

Examples 2.2. **(1) The rank-one case.** In case $N = 1$, the only choice of R is $R = \{\pm\sqrt{2}\}$, which is the root system of type A_1. The corresponding reflection group is $G = \{id, \sigma\}$ acting on $\mathbb{R}$ by $\sigma(x) = -x$. The Dunkl operator $T := T_1$ associated with the multiplicity parameter $k \in \mathbb{C}$ is given by

$$Tf(x) \ = \ f'(x) + k\, \frac{f(x) - f(-x)}{x}.$$

Its square T^2, when restricted to the even subspace $C^2(\mathbb{R})^e := \{f \in C^2(\mathbb{R}) : f(x) = f(-x)\}$, is given by a singular Sturm-Liouville operator:

$$T^2|_{C^2(\mathbb{R})^e} f \;=\; f'' + \frac{2k}{x}\cdot f'\,.$$

(2) Dunkl operators of type A_{N-1} **.** Suppose $G = S_N$ with root system of type A_{N-1}. (In contrast to the above example, G now acts on $\mathbb{R}^N$). As all transpositions are conjugate in S_N, the vector space of multiplicity functions is one-dimensional. The Dunkl operators associated with the multiplicity parameter $k \in \mathbb{C}$ are given by

$$T_i^S \;=\; \partial_i \,+\, k\cdot\sum_{j\neq i}\frac{1-\sigma_{ij}}{x_i - x_j} \quad (i = 1,\ldots,N),$$

and the k-Laplacian or Dunkl Laplacian is

$$\Delta_k^S \;=\; \Delta + 2k \sum_{1\leq i<j\leq N}\frac{1}{x_i - x_j}\big[(\partial_i - \partial_j) - \frac{1-\sigma_{ij}}{x_i - x_j}\big]\,.$$

(3) Dunkl operators of type B_N **.** Suppose R is a root system of type B_N, corresponding to $G = S_N \ltimes \mathbb{Z}_2^N$. There are two conjugacy classes of reflections in G, leading to multiplicity functions of the form $k = (k_0, k_1)$ with $k_i \in \mathbb{C}$. The associated Dunkl operators are given by

$$T_i^B \;=\; \partial_i \,+\, k_1\frac{1-\sigma_i}{x_i} \,+\, k_0\cdot\sum_{j\neq i}\big[\frac{1-\sigma_{ij}}{x_i - x_j} + \frac{1-\tau_{ij}}{x_i + x_j}\big] \quad (i = 1,\ldots,N),$$

where $\tau_{ij} := \sigma_{ij}\sigma_i\sigma_j$.

2.3 A formula of Macdonald and its analog in Dunkl theory

In the classical theory of spherical harmonics (see for instance [28]) the following bilinear pairing on Π, sometimes called Fischer product, plays an important role:

$$[p,q]_0 := (p(\partial)q)(0), \quad p,q \in \Pi.$$

This pairing is closely related to the scalar product in $L^2(\mathbb{R}^N, e^{-|x|^2/2}dx)$; in fact, in his short note [39] Macdonald observed the following identity:

$$[p,q]_0 \;=\; (2\pi)^{-N/2}\int_{\mathbb{R}^N} e^{-\Delta/2}p(x)\, e^{-\Delta/2}q(x)\, e^{-|x|^2/2}\, dx.$$

Here $e^{-\Delta/2}$ is well-defined as a linear operator on Π by means of the terminating series

$$e^{-\Delta/2}p = \sum_{n=0}^{\infty}\frac{(-1)^n}{2^n n!}\Delta^n p.$$

Both the Fischer product as well as Macdonald's identity have a useful generalization in the Dunkl setting. In the following, we shall always restrict to the case $k \geq 0$.

Definition 2.4. *For $p, q \in \Pi$ define*

$$[p, q]_k := (p(T)q)(0).$$

This bilinear form was introduced in [14]. We collect some of its basic properties:

Lemma 2.4. (1) *If $p \in \mathcal{P}_n$ and $q \in \mathcal{P}_m$ with $n \neq m$, then $[p, q]_k = 0$.*
(2) $[x_i\, p\,, q]_k = [p, T_i\, q]_k \quad (p, q \in \Pi,\ i = 1, \ldots, N)$.
(3) $[g \cdot p\,, g \cdot q]_k = [p, q]_k \quad (p, q \in \Pi,\ g \in G)$.

Proof. (1) follows from the homogeneity of the Dunkl operators, (2) is clear from the definition, and (3) follows from Exercise 3.

Let w_k denote the weight function on $\mathbb{R}^N$ defined by

$$w_k(x) = \prod_{\alpha \in R_+} |\langle \alpha, x \rangle|^{2k(\alpha)}. \tag{2.3}$$

It is G-invariant and homogeneous of degree 2γ, with the index

$$\gamma := \gamma(k) := \sum_{\alpha \in R_+} k(\alpha). \tag{2.4}$$

Notice that by G-invariance of k, we have $k(-\alpha) = k(\alpha)$ for all $\alpha \in R$. Hence this definition does again not depend on the special choice of R_+. Further, we define the constant

$$c_k := \int_{\mathbb{R}^N} e^{-|x|^2/2} w_k(x) dx,$$

a a so-called Macdonald-Mehta-Selberg integral. There exists a closed form for it which was conjectured and proved by Macdonald [40] for the infinite series of root systems. An extension to arbitrary crystallographic reflection groups is due to Opdam [44], and there are computer-assisted proofs for some non-crystallographic root systems. As far as we know, a general proof for arbitrary root systems has not yet been found.

We shall need the following anti-symmetry of the Dunkl operators:

Proposition 2.1. [15] *Let $k \geq 0$. Then for every $f \in \mathscr{S}(\mathbb{R}^N)$ and $g \in C_b^1(\mathbb{R}^N)$,*

$$\int_{\mathbb{R}^N} T_\xi f(x) g(x) w_k(x) dx = -\int_{\mathbb{R}^N} f(x) T_\xi g(x) w_k(x) dx\,.$$

Proof. A short calculation. In order to have the appearing integrals well defined, one has to assume $k \geq 1$ first, and then extend the result to general $k \geq 0$ by analytic continuation.

Proposition 2.2. *For all* $p, q \in \Pi$,

$$[p,q]_k \,=\, c_k^{-1}\int_{\mathbb{R}^N} e^{-\Delta_k/2}p(x)\, e^{-\Delta_k/2}q(x)\, e^{-|x|^2/2}\, w_k(x)dx. \qquad (2.5)$$

This result is due to Dunkl ([14]). As the Dunkl Laplacian is homogeneous of degree -2, the operator $e^{-\Delta_k/2}$ is well-defined and bijective on Π, and it preserves the degree. We give here a direct proof which is partly taken from an unpublished part of M. de Jeu's thesis ([31], §3.3). It involves the following commutator results in $\mathrm{End}_{\mathbb{C}}(\Pi)$, where as usual, $[A,B] = AB - BA$ for $A, B \in \mathrm{End}_{\mathbb{C}}(\Pi)$.

Lemma 2.5. *For* $i = 1, \ldots, N$,

(1) $[x_i\,,\, \Delta_k/2] \,=\, -T_i\,$;
(2) $[x_i\,,\, e^{-\Delta_k/2}] \,=\, T_i\, e^{-\Delta_k/2}$.

Proof. (1) follows by direct calculation, c.f. [12]. Induction then yields that

$$[x_i\,, (\Delta_k/2)^n] \,=\, -n\, T_i(\Delta_k/2)^{n-1} \quad \text{for } n \geq 1,$$

and this implies (2).

Proof (of Proposition 2.2). Let $i \in \{1, \ldots, N\}$, and denote the right side of (2.5) by $(p,q)_k$. Then by the anti-symmetry of T_i in $L^2(\mathbb{R}^N, w_k)$, the product rule for T_i, see Exercise 4, and the above Lemma,

$$\begin{aligned}
(p, T_i\, q)_k \,&=\, c_k^{-1}\int_{\mathbb{R}^N} e^{-\Delta_k/2}p\cdot (T_i\, e^{-\Delta_k/2}q)\, e^{-|x|^2/2}\, w_k\, dx \\
&=\, -\,c_k^{-1}\int_{\mathbb{R}^N} T_i(e^{-|x|^2/2}e^{-\Delta_k/2}p)\cdot (e^{-\Delta_k/2}q)\, w_k\, dx \\
&=\, c_k^{-1}\int_{\mathbb{R}^N} e^{-\Delta_k/2}(x_i\, p)\cdot (e^{-\Delta_k/2}q)\, e^{-|x|^2/2}\, w_k\, dx \,=\, (x_i\, p, q)_k.
\end{aligned}$$

But the form $[\,.,.\,]_k$ has the same property by Lemma 2.4(2). It is now easily checked that the assertion is true if p or q is constant, and then, by induction on $\max(\deg p, \deg q)$, for all homogeneous p, q. This suffices by the linearity of both forms.

Corollary 2.1. *Let again* $k \geq 0$. *Then the pairing* $[\,.\,,\,.\,]_k$ *on* Π *is symmetric and non-degenerate, i.e.* $[p,q]_k = 0$ *for all* $q \in \Pi$ *implies that* $p = 0$.

Exercise 5. Check the details in the proofs of Proposition 2.2 and Corollary 2.1.

2.4 Dunkl's intertwining operator

It was first shown in [14] that for non-negative multiplicity functions, the associated commutative algebra of Dunkl operators is intertwined with the algebra of usual partial differential operators by a unique linear and homogeneous isomorphism on polynomials. A thorough analysis in [18] subsequently revealed that for general k, such an intertwining operator exists if and only if the common kernel of the T_ξ, considered as linear operators on Π, contains no "singular" polynomials besides the constants. More precisely, the following characterization holds:

Theorem 2.3. [18] *Let* $K^{reg} := \{k \in K : \bigcap_{\xi\in\mathbb{R}^N} Ker T_\xi(k) = \mathbb{C}\cdot 1\}$. *Then the following statements are equivalent:*

(1) $k \in K^{reg}$;
(2) *There exists a unique linear isomorphism ("intertwining operator")* V_k *of* Π *such that*

$$V_k(\mathcal{P}_n) = \mathcal{P}_n\,,\ \ V_k|_{\mathcal{P}_0} = id \quad \text{and} \ \ T_\xi V_k = V_k\,\partial_\xi \quad \text{for all } \xi \in \mathbb{R}^N.$$

The proof of this result is by induction on the degree of homogeneity and requires only linear algebra.

The intertwining operator V_k commutes with the G-action:

Exercise 6. $g^{-1}\circ V_k \circ g = V_k \quad (g \in G)$.
Hint: Use the G-equivariance of the T_ξ and the defining properties of V_k.

Proposition 2.3. $\{k \in K : k \geq 0\} \subseteq K^{reg}$.

Proof. Suppose that $p \in \oplus_{n\geq 1}\mathcal{P}_n$ satisfies $T_\xi(k)p = 0$ for all $\xi \in \mathbb{R}^N$. Then $[q,p]_k = 0$ for all $q \in \oplus_{n\geq 1}\mathcal{P}_n$, and hence also $[q,p]_k = 0$ for all $q \in \Pi$. Thus $p = 0$, by the non-degeneracy of $[\,.\,,\,.\,]_k$, see Corollary 2.1.

The complete singular parameter set $K \setminus K^{reg}$ is explicitly determined in [18]. It is an open subset of K which is invariant under complex conjugation, and contains $\{k \in K : \text{Re}\, k \geq 0\}$. Later in these lectures, we will in fact restrict our attention to non-negative multiplicity functions. These are of particular interest concerning our subsequent positivity results, which could not be expected for non-positive multiplicities. Though the intertwining operator plays an important role in Dunkl's theory, an explicit "closed" form for it is known so far only in some special cases. Among these are

1. *The rank-one case.* Here

$$K^{reg} = \mathbb{C}\setminus\{-1/2 - n,\ n\in\mathbb{Z}_+\}.$$

The associated intertwining operator is given explicitly by

$$V_k(x^{2n}) \,=\, \frac{(\frac{1}{2})_n}{(k+\frac{1}{2})_n}\, x^{2n}\,; \quad V_k(x^{2n+1}) \,=\, \frac{(\frac{1}{2})_{n+1}}{(k+\frac{1}{2})_{n+1}}\, x^{2n+1}\,,$$

where $(a)_n = \Gamma(a+n)/\Gamma(a)$ is the Pochhammer-symbol. For $\mathrm{Re}\, k > 0$, this amounts to the following integral representation (see [14], Th. 5.1):

$$V_k\, p(x) \,=\, \frac{\Gamma(k+1/2)}{\Gamma(1/2)\,\Gamma(k)} \int_{-1}^{1} p(xt)\,(1-t)^{k-1}(1+t)^k\, dt. \tag{2.6}$$

2. *The case* $G = S_3$. This was studied in [16]. Here

$$K^{reg} = \mathbb{C} \setminus \{-1/2 - n, -1/3 - n, -2/3 - n,\, n \in \mathbb{Z}_+\}.$$

In order to bring V_k into action in a further development of the theory, it is important to extend it to larger function spaces. For this we shall always assume that $k \geq 0$. In a first step, V_k is extended to a bounded linear operator on suitably normed algebras of homogeneous series on a ball. This concept goes back to [14].

Definition 2.5. *For* $r > 0$, *let* $B_r := \{x \in \mathbb{R}^N : |x| \leq r\}$ *denote the closed ball of radius* r, *and let* A_r *be the closure of* Π *with respect to the norm*

$$\|p\|_{A_r} := \sum_{n=0}^{\infty} \|p_n\|_{\infty, B_r} \quad \textit{for } p = \sum_{n=0}^{\infty} p_n\,,\, p_n \in \mathcal{P}_n.$$

Clearly A_r is a commutative Banach-$*$-algebra under the pointwise operations and with complex conjugation as involution. Each $f \in A_r$ has a unique representation $f = \sum_{n=0}^{\infty} f_n$ with $f_n \in \mathcal{P}_n$, and is continuous on the ball B_r and real-analytic in its interior. The topology of A_r is stronger than the topology induced by the uniform norm on B_r. Notice also that A_r is not closed with respect to $\|.\|_{\infty,B_r}$ and that $A_r \subseteq A_s$ with $\|\,.\,\|_{A_r} \geq \|\,.\,\|_{A_s}$ for $s \leq r$.

Theorem 2.4. $\|V_k p\|_{\infty,B_r} \leq \|p\|_{\infty,B_r}$ *for each* $p \in \mathcal{P}_n$.

The proof of this result is given in [14] and can also be found in [19]. It uses the *van der Corput-Schaake inequality* which states that for each *real-valued* $p \in \mathcal{P}_n$,

$$\sup\{|\langle \nabla p(x), y\rangle| : x, y \in B_1\} \,\leq\, n\|p\|_{\infty,B_1}.$$

Notice that here the converse inequality is trivially satisfied, because for $p \in \mathcal{P}_n$ we have $\langle \nabla p(x), x\rangle = np(x)$. The following is now immediate:

Corollary 2.2. $\|V_k f\|_{A_r} \leq \|f\|_{A_r}$ *for every* $f \in \Pi$, *and* V_k *extends uniquely to a bounded linear operator on* A_r *via*

$$V_k f := \sum_{n=0}^{\infty} V_k f_n \quad \textit{for } f = \sum_{n=0}^{\infty} f_n\,.$$

Formula (2.6) shows in particular that in the rank-one case with $k > 0$, the operator V_k is positivity-preserving on polynomials. It was conjectured by Dunkl in [14] that for arbitrary reflection groups and non-negative multiplicity functions, the linear functional $f \mapsto V_k f(x)$ on A_r should be positive. We shall see in Section 4.1 that this is in fact true. As a consequence, we shall obtain the existence of a positive integral representation generalizing (2.6), which in turn allows to extend V_k to larger function spaces. This positivity result also has important consequences for the structure of the Dunkl kernel, which generalizes the usual exponential function in the Dunkl setting. We shall introduce it in the following section.

Exercise 7. The *symmetric spectrum* $\Delta_S(A)$ of a (unital) commutative Banach-$*$-algebra A is defined as the set of all non-zero algebra homomorphisms $\varphi : A \to \mathbb{C}$ satisfying the $*$-condition $\varphi(a^*) = \overline{\varphi(a)}$ for all $a \in A$. It is a compact Hausdorff space with the weak-*-topology (sometimes called the Gelfand topology). Prove that the symmetric spectrum of the algebra A_r is given by $\Delta_S(A_r) = \{\varphi_x : x \in B_r\}$, where φ_x is the evaluation homomorphism $\varphi_x(f) := f(x)$. Show also that the mapping $x \mapsto \varphi_x$ is a homeomorphism from B_r onto $\Delta_S(A_r)$.

2.5 The Dunkl kernel

Throughout this section we assume that $k \geq 0$. Moreover, we denote by $\langle\,.\,,\,.\,\rangle$ not only the Euclidean scalar product on $\mathbb{R}^N$, but also its *bilinear* extension to $\mathbb{C}^N \times \mathbb{C}^N$. For fixed $y \in \mathbb{C}^N$, the exponential function $x \mapsto e^{\langle x,y\rangle}$ belongs to each of the algebras A_r, $r > 0$. This justifies the following

Definition 2.6. [14] *For $y \in \mathbb{C}^N$, define*

$$E_k(x,y) := V_k(e^{\langle\,.\,,y\rangle})(x), \quad x \in \mathbb{R}^N.$$

The function E_k is called the *Dunkl-kernel*, or k-exponential kernel, associated with G and k. It can alternatively be characterized as the solution of a joint eigenvalue problem for the associated Dunkl operators.

Proposition 2.4. *Let $k \geq 0$ and $y \in \mathbb{C}^N$. Then $f = E_k(\,.\,,y)$ is the unique solution of the system*

$$T_\xi f = \langle \xi, y\rangle f \quad \text{for all } \xi \in \mathbb{R}^N \tag{2.7}$$

which is real-analytic on $\mathbb{R}^N$ and satisfies $f(0) = 1$.

Proof. $E_k(\,.\,y)$ is real-analytic on $\mathbb{R}^N$ by our construction. Define

$$E_k^{(n)}(x,y) := \frac{1}{n!} V_k \langle\,.\,,y\rangle^n(x), \quad x \in \mathbb{R}^N,\, n = 0,1,2,\ldots.$$

Then $E_k(x,y) = \sum_{n=0}^\infty E_k^{(n)}(x,y)$, and the series converges uniformly and absolutely with respect to x. The homogeneity of V_k immediately implies $E_k(0,y) = 1$. Further, by the intertwining property,

$$T_\xi E_k^{(n)}(\,.\,,y) \,=\, \frac{1}{n!} V_k\, \partial_\xi \langle\,.\,,y\rangle^n \,=\, \langle \xi, y\rangle E_k^{(n-1)}(\,.\,,y) \tag{2.8}$$

for all $n \geq 1$. This shows that $E_k(\,.\,,y)$ solves (2.7). To prove uniqueness, suppose that f is a real-analytic solution of (2.7) with $f(0) = 1$. Then T_ξ can be applied termwise to the homogeneous expansion $f = \sum_{n=0}^\infty f_n\,, f_n \in \mathcal{P}_n$, and comparison of homogeneous parts shows that

$$f_0 = 1, \quad T_\xi f_n = \langle \xi, y\rangle f_{n-1} \quad \text{for } n \geq 1.$$

As $\{k \in K : k \geq 0\} \subseteq K^{reg}$, it follows by induction that all f_n are uniquely determined.

While this construction has been carried out only for $k \geq 0$, there is a more general result by Opdam which assures the existence of a general exponential kernel with properties according to the above lemma for arbitrary regular multiplicity parameters. The following is a weakened version of [44], Proposition 6.7; it in particular implies that E_k has a holomorphic extension to $\mathbb{C}^N \times \mathbb{C}^N$:

Theorem 2.5. *For each $k \in K^{reg}$ and $y \in \mathbb{C}^N$, the system*

$$T_\xi\, f \,=\, \langle \xi, y\rangle f \quad (\xi \in \mathbb{R}^N)$$

has a unique solution $x \mapsto E_k(x,y)$ which is real-analytic on $\mathbb{R}^N$ and satisfies $f(0) = 1$. Moreover, the mapping $(x,k,y) \mapsto E_k(x,y)$ extends to a meromorphic function on $\mathbb{C}^N \times K \times \mathbb{C}^N$ with pole set $\mathbb{C}^N \times (K \setminus K^{reg}) \times \mathbb{C}^N$

We collect some further properties of the Dunkl kernel E_k.

Proposition 2.5. *Let $k \geq 0$, $x, y \in \mathbb{C}^N$, $\lambda \in \mathbb{C}$ and $g \in G$.*

(1) $E_k(x,y) = E_k(y,x)$
(2) $E_k(\lambda x, y) = E_k(x, \lambda y)$ *and* $E_k(gx, gy) = E_k(x,y)$.
(3) $\overline{E_k(x,y)} = E_k(\overline{x}, \overline{y})$.

Proof. (1) This is shown in [14]. (2) is easily obtained from the definition of E_k together with the homogeneity and equivariance properties of V_k. For (3), notice that $f := \overline{E_k(\,.\,,y)}$, which is again real-analytic on $\mathbb{R}^N$, satisfies $T_\xi f \,=\, \langle \xi, \overline{y}\rangle\, f$, $f(0) = 1$. By the uniqueness part of the above Proposition, $\overline{E_k(x,y)} = E_k(x, \overline{y})$ for all real x. Now both $x \mapsto \overline{E_k(\overline{x},y)}$ and $x \mapsto E_k(x,\overline{y})$ are holomorphic on $\mathbb{C}^N$ and agree on $\mathbb{R}^N$. Hence they coincide.

Just as with the intertwining operator, the kernel E_k is explicitly known for some particular cases only. An important example is again the rank-one situation:

Example 2.1. In the *rank-one case* with $\operatorname{Re} k > 0$, the integral representation (2.6) for V_k implies that for all $x, y \in \mathbb{C}$,

$$\begin{aligned} E_k(x,y) &= \frac{\Gamma(k+1/2)}{\Gamma(1/2)\,\Gamma(k)} \int_{-1}^{1} e^{txy}(1-t)^{k-1}(1+t)^k\,dt \\ &= e^{xy}\cdot {}_1F_1(k, 2k+1, -2xy). \end{aligned}$$

This can also be written as

$$E_k(x,y) = j_{k-1/2}(ixy) + \frac{xy}{2k+1}\, j_{k+1/2}(ixy),$$

where for $\alpha \geq -1/2$, j_α is the normalized spherical Bessel function

$$j_\alpha(z) = 2^\alpha \Gamma(\alpha+1)\cdot \frac{J_\alpha(z)}{z^\alpha} = \Gamma(\alpha+1)\cdot \sum_{n=0}^{\infty} \frac{(-1)^n (z/2)^{2n}}{n!\,\Gamma(n+\alpha+1)}\,. \tag{2.9}$$

This motivates the following

Definition 2.7. [44] *The k-Bessel function (or generalized Bessel function) is defined by*

$$J_k(x,y) := \frac{1}{|G|}\sum_{g\in G} E_k(gx, y) \quad (x, y \in \mathbb{C}^N). \tag{2.10}$$

Thanks to Proposition 2.5 J_k is G-invariant in both arguments and therefore naturally considered on Weyl chambers of G (or their complexifications). In the rank-one case, we have

$$J_k(x,y) = j_{k-1/2}(ixy).$$

It is a well-known fact from classical analysis that for fixed $y \in \mathbb{C}$, the function $f(x) = j_{k-1/2}(ixy)$ is the unique analytic solution of the differential equation

$$f'' + \frac{2k}{x} f' = y^2 f$$

which is even and normalized by $f(0) = 1$. In order to see how this can be generalized to the multivariable case, consider the algebra of G-invariant polynomials on $\mathbb{R}^N$,

$$\Pi^G = \{p \in \Pi : g\cdot p = p \quad \text{for all } g \in G\}.$$

If $p \in \Pi^G$, then it follows from the equivariance of the Dunkl operators (Exercise 3) that $p(T)$ commutes with the G-action; a detailed argument for this is given in [26]. Thus $p(T)$ leaves Π^G invariant, and we obtain in particular that for fixed $y \in \mathbb{C}^N$, the k-Bessel function $J_k(\,.\,, y)$ is a solution to the following Bessel-system:

$$p(T)f = p(y)f \quad \text{for all } p \in \Pi^G, \ f(0) = 1.$$

According to [44], it is in fact the only G-invariant and analytic solution. We mention that there exists a group theoretic context in which, for a certain parameters k, generalized Bessel functions occur in a natural way: namely as the spherical functions of a Euclidean type symmetric space, associated with a so-called Cartan motion group. We refer to [44] for this connection and to [28] for the necessary background in semisimple Lie theory.

The Dunkl kernel is of particular interest as it gives rise to an associated integral transform on $\mathbb{R}^N$ which generalizes the Euclidean Fourier transform in a natural way. This transform will be discussed in the following section. Its definition and essential properties rely on suitable growth estimates for E_k. In our case $k \geq 0$, the best ones to be expected are available:

Proposition 2.6. [51] *For all $x \in \mathbb{R}^N$, $y \in \mathbb{C}^N$ and all multi-indices $\alpha \in \mathbb{Z}_+^N$,*

$$|\partial_y^\alpha E_k(x,y)| \leq |x|^{|\alpha|} \max_{g \in G} e^{Re\langle gx, y\rangle}.$$

In particular, $|E_k(-ix,y)| \leq 1$ *for all* $x, y \in \mathbb{R}^N$.

This result will be obtained later from a positive integral representation of Bochner-type for E_k, c.f. Corollary 4.1. M. de Jeu had slightly weaker bounds in [30], differing by an additional factor $\sqrt{|G|}$.

We conclude this section with two important reproducing properties for the Dunkl kernel. Notice that the above estimate on E_k assures the convergence of the involved integrals.

Proposition 2.7. *Let $k \geq 0$. Then for $p \in \Pi$, $y, z \in \mathbb{C}^N$,*

(1) $\displaystyle\int_{\mathbb{R}^N} e^{-\Delta_k/2} p(x)\, E_k(x,y)\, e^{-|x|^2/2} w_k(x)dx = c_k\, e^{\langle y,y\rangle/2} p(y),$

(2) $\displaystyle\int_{\mathbb{R}^N} E_k(x,y)\, E_k(x,z)\, e^{-|x|^2/2} w_k(x)dx = c_k\, e^{(\langle y,y\rangle + \langle z,z\rangle)/2} E_k(y,z).$

Proof. (c.f. [15].) We shall use the Macdonald-type formula (2.5) for the pairing $[\,.\,,\,.\,]_k$. First, we prove that

$$[E_k^{(n)}(x,\,.\,),\,.\,]_k = p(x) \quad \text{for all } p \in \mathcal{P}_n,\ x \in \mathbb{R}^N. \tag{2.11}$$

In fact, if $p \in \mathcal{P}_n$, then

$$p(x) = (\langle x, \partial_y\rangle^n/n!)p(y) \quad \text{and} \quad V_k^x p(x) = E_k^{(n)}(x, \partial_y)p(y).$$

Here the uppercase index in V_k^x denotes the relevant variable. Application of V_k^y to both sides yields $V_k^x p(x) = E_k^{(n)}(x, T^y) V_k^y p(y)$. As V_k is bijective on $\mathcal{P}_n$, this implies (2.11). For fixed y, let $L_n(x) := \sum_{j=0}^n E_k^{(j)}(x,y)$. If n is

larger than the degree of p, it follows from (2.11) that $[L_n, p]_k = p(y)$. Thus in view of the Macdonald formula,

$$c_k^{-1} \int_{\mathbb{R}^N} e^{-\Delta_k/2} L_n(x) e^{-\Delta_k/2} p(x) e^{-|x|^2/2} w_k(x) dx = p(y).$$

On the other hand, it is easily checked that

$$\lim_{n\to\infty} e^{-\Delta_k/2} L_n(x) = e^{-\langle y,y\rangle/2} E_k(x,y).$$

This gives (1). Identity (2) then follows from (1), again by homogeneous expansion of E_k.

2.6 The Dunkl transform

The Dunkl transform was introduced in [15] for non-negative multiplicity functions and further studied in [30] in the more general case $\operatorname{Re} k \geq 0$. In these notes, we again restrict ourselves to $k \geq 0$.

Definition 2.8. *The Dunkl transform associated with G and $k \geq 0$ is given by*

$$\widehat{\cdot}^{\,k} : L^1(\mathbb{R}^N, w_k) \to C_b(\mathbb{R}^N);$$
$$\widehat{f}^{\,k}(\xi) := c_k^{-1} \int_{\mathbb{R}^N} f(x) E_k(-i\xi, x)\, w_k(x) dx \qquad (\xi \in \mathbb{R}^N).$$

The inverse transform is defined by $f^{\vee k}(\xi) = \widehat{f}^{\,k}(-\xi)$.

Notice that $\widehat{f}^{\,k} \in C_b(\mathbb{R}^N)$ results from our bounds on E_k. The Dunkl transform shares many properties with the classical Fourier transform. Here are the most basic ones:

Lemma 2.6. *Let $f \in \mathscr{S}(\mathbb{R}^N)$. Then for $j = 1, \ldots, N$,*

(1) $\widehat{f}^{\,k} \in C^\infty(\mathbb{R}^N)$ *and* $T_j(\widehat{f}^{\,k}) = -(ix_j f)^{\wedge k}$.
(2) $(T_j f)^{\wedge k}(\xi) = i\xi_j \widehat{f}^{\,k}(\xi)$.
(3) *The Dunkl transform leaves $\mathscr{S}(\mathbb{R}^N)$ invariant.*

Proof. (1) is obvious from (2.7), and (2) follows from the anti-symmetry relation (Proposition 2.1) for the Dunkl operators. For (3), notice that it suffices to prove that $\partial_\xi^\alpha(\xi^\beta \widehat{f}^{\,k}(\xi))$ is bounded for arbitrary multi-indices α, β. By the previous Lemma, we have $\xi^\beta \widehat{f}^{\,k}(\xi) = \widehat{g}^{\,k}(\xi)$ for some $g \in \mathscr{S}(\mathbb{R}^N)$. Using the growth bounds of Proposition 2.6 yields the assertion.

Exercise 8.

(1) $C_c^\infty(\mathbb{R}^N)$ and $\mathscr{S}(\mathbb{R}^N)$ are dense in $L^p(\mathbb{R}^N, w_k),\ p = 1, 2.$
(2) Conclude the *Lemma of Riemann-Lebesgue* for the Dunkl transform:

$$f \in L^1(\mathbb{R}^N, w_k) \implies \widehat{f}^{\,k} \in C_0(\mathbb{R}^N).$$

Here $C_0(\mathbb{R}^N)$ denotes the space of continuous functions on $\mathbb{R}^N$ which vanish at infinity.

The following are the main results for the Dunkl transform; we omit the proofs but refer the reader to [15] and [30]:

Theorem 2.6. (1) *The Dunkl transform* $f \mapsto \widehat{f}^{\,k}$ *is a homeomorphism of* $\mathscr{S}(\mathbb{R}^N)$ *with period* 4.
(2) *(Plancherel theorem) The Dunkl transform has a unique extension to an isometric isomorphism of* $L^2(\mathbb{R}^N, w_k)$. *We denote this isomorphism again by* $f \mapsto \widehat{f}^{\,k}$.
(3) *(L^1-inversion) For all* $f \in L^1(\mathbb{R}^N, w_k)$ *with* $\widehat{f}^{\,k} \in L^1(\mathbb{R}^N, w_k)$,

$$f = (\widehat{f}^{\,k})^{\vee k} \quad a.e.$$

3 CMS models and generalized Hermite polynomials

3.1 Quantum Calogero-Moser-Sutherland models

Quantum Calogero-Moser-Sutherland (CMS) models describe quantum mechanical systems of N identical particles on a circle or line which interact pairwise through long range potentials of inverse square type. They are exactly solvable and have gained considerable interest in theoretical physics during the last years. Among the broad literature in this area, we refer to [10], [36], [33], [5], [2]-[4], [46], [47], [61], [17]. CMS models have in particular attracted some attention in conformal field theory, and they are being used to test the ideas of fractional statistics ([24], [25]). While explicit spectral resolutions of such models were already obtained by Calogero and Sutherland ([6], [57]), a new aspect in the understanding of their algebraic structure and quantum integrability was much later initiated by [48] and [26]. The Hamiltonian under consideration is hereby modified by certain exchange operators, which allow to write it in a decoupled form. These exchange modifications can be expressed in terms of Dunkl operators of type A_{N-1}. The Hamiltonian of the *linear CMS model with harmonic confinement* in $L^2(\mathbb{R}^N)$ is given by

$$\mathcal{H}_C = -\Delta + g \sum_{1 \le i < j \le N} \frac{1}{(x_i - x_j)^2} + \omega^2 |x|^2; \tag{3.1}$$

here $\omega > 0$ is a frequency parameter and $g \geq -1/2$ is a coupling constant. In case $\omega = 0$, (3.1) describes the free Calogero model. On the other hand, if $g = 0$, then $\mathcal{H}_C$ coincides with the Hamiltonian of the N-dimensional isotropic harmonic oscillator,

$$\mathcal{H}_0 = -\Delta + \omega^2|x|^2.$$

The spectral decomposition of this operator in $L^2(\mathbb{R}^N)$ is well-known: The spectrum is discrete, $\sigma(\mathcal{H}_0) = \{(2n+N)\omega,\, n \in \mathbb{Z}_+\}$, and the classical multivariable Hermite functions (tensor products of one-variable Hermite functions, c.f. Examples 3.1), form a complete set of eigenfunctions. The study of the Hamiltonian $\mathcal{H}_C$ was initiated by Calogero ([6]); he computed its spectrum and determined the structure of the bosonic eigenfunctions and scattering states in the confined and free case, respectively. Perelomov [47] observed that (3.1) is completely quantum integrable, i.e. there exist N commuting, algebraically independent symmetric linear operators in $L^2(\mathbb{R}^N)$ including $\mathcal{H}_C$. We mention that the complete integrability of the classical Hamiltonian systems associated with (3.1) goes back to Moser [41]. There exist generalizations of the classical Calogero-Moser-Sutherland models in the context of abstract root systems, see for instance [42], [43]. In particular, if R is an arbitrary root system on $\mathbb{R}^N$ and k is a nonnegative multiplicity function on it, then the corresponding abstract Calogero Hamiltonian with harmonic confinement is given by

$$\widetilde{\mathcal{H}}_k \,=\, -\widetilde{\mathcal{F}}_k + \omega^2|x|^2$$

with the formal expression

$$\widetilde{\mathcal{F}}_k \,=\, \Delta - \,2\sum_{\alpha\in R_+} k(\alpha)(k(\alpha)-1)\frac{1}{\langle\alpha,x\rangle^2}.$$

If R is of type A_{N-1}, then $\widetilde{\mathcal{H}}_k$ just coincides with $\mathcal{H}_C$. For both the classical and the quantum case, partial results on the integrability of this model are due to Olshanetsky and Perelomov [42], [43]. A new aspect in the understanding of the algebraic structure and the quantum integrability of CMS systems was initiated by Polychronakos [48] and Heckman [26]. The underlying idea is to construct quantum integrals for CMS models from differential-reflection operators. Polychronakos introduced them in terms of an "exchange-operator formalism" for (3.1). He thus obtained a complete set of commuting observables for (3.1) in an elegant way. In [26] it was observed in general that the complete algebra of quantum integrals for free, abstract Calogero models is intimately connected with the corresponding algebra of Dunkl operators. Let us briefly describe this connection: Consider the following modification of $\widetilde{\mathcal{F}}_k$, involving reflection terms:

$$\mathcal{F}_k \,=\, \Delta - 2\sum_{\alpha\in R_+}\frac{k(\alpha)}{\langle\alpha,x\rangle^2}\,(k(\alpha)-\sigma_\alpha)\,. \tag{3.2}$$

In order to avoid singularities in the reflecting hyperplanes, it is suitable to carry out a gauge transform by $w_k^{1/2}$. A short calculation, using again results from [12], gives

$$w_k^{-1/2}\mathcal{F}_k w_k^{1/2} \,=\, \Delta_k\,,$$

c.f. [52]. Here Δ_k is the Dunkl Laplacian associated with G and k. Now consider the algebra of Π^G of G-invariant polynomials on $\mathbb{R}^N$. By a classical theorem of Chevalley (see e.g. [29]), it is generated by N homogeneous, algebraically independent elements. For $p \in \Pi^G$ we denote by $\operatorname{Res}(p(T))$ the restriction of the Dunkl operator $p(T)$ to Π^G (Recall that $p(T)$ leaves Π^G invariant!). Then

$$\mathcal{A} := \{\operatorname{Res} p(T) :\, p \in \Pi^G\}$$

is a commutative algebra of differential operators on Π^G containing the operator

$$\operatorname{Res}(\Delta_k) \,=\, w_k^{-1/2}\widetilde{\mathcal{F}}_k\, w_k^{1/2},$$

and $\mathcal{A}$ has N algebraically independent generators, called quantum integrals for the free Hamiltonian $\widetilde{\mathcal{F}}_k$.

3.2 Spectral analysis of abstract CMS Hamiltonians

This section is devoted to a spectral analysis of abstract linear CMS operators with harmonic confinement. We follow the expositions in [50], [53]. To simplify formulas, we fix $\omega = 1/2$; corresponding results for general ω can always be obtained by rescaling. We again work with the gauge-transformed version with reflection terms,

$$\mathcal{H}_k := \, w_k^{-1/2}(-\mathcal{F}_k + \frac{1}{4}|x|^2)\, w_k^{1/2} \,=\, -\Delta_k + \frac{1}{4}|x|^2.$$

Due to the anti-symmetry of the first order Dunkl operators (Proposition 2.1), this operator is symmetric and densely defined in $L^2(\mathbb{R}^N, w_k)$ with domain $\mathcal{D}(\mathcal{H}_k) := \mathscr{S}(\mathbb{R}^N)$. Notice that in case $k = 0$, $\mathcal{H}_k$ is just the Hamiltonian of the N-dimensional isotropic harmonic oscillator. We further consider the Hilbert space $L^2(\mathbb{R}^N, m_k)$, where m_k is the probability measure

$$dm_k \,:=\, c_k^{-1}e^{-|x|^2/2}w_k(x)dx \tag{3.3}$$

and the operator

$$\mathcal{J}_k := \, -\Delta_k + \sum_{i=1}^N x_i\partial_i$$

in $L^2(\mathbb{R}^N, m_k)$, with domain $\mathcal{D}(\mathcal{J}_k) := \Pi$. It can be shown by standard methods that Π is dense in $L^2(\mathbb{R}^N, m_k)$. We do not carry this out; a proof can be found in [51] or in [32], where a comprehensive treatment of density questions in several variables is given.

The next theorem contains a complete description of the spectral properties of $\mathcal{H}_k$ and $\mathcal{J}_k$ and generalizes the already mentioned well-known facts for the classical harmonic oscillator Hamilonian. For the proof, we shall employ the $sl(2)$-commutation relations of the operators

$$E := \frac{1}{2}|x|^2, \;\; F := -\frac{1}{2}\Delta_k \quad \text{and} \;\; H := \sum_{i=1}^N x_i\partial_i + (\gamma + N/2)$$

on Π (with the index $\gamma = \gamma(k)$ as defined in (2.4)) which can be found in [26]. They are

$$[H, E] = 2E, \quad [H, F] = -2F, \quad [E, F] = H. \tag{3.4}$$

Notice that the first two relations are immediate consequences of the fact that the Euler operator

$$\rho := \sum_{i=1}^N x_i\partial_i \tag{3.5}$$

satisfies $\rho(p) = np$ for each homogeneous $p \in \mathcal{P}_n$. We start with the following

Lemma 3.1. *On* $\mathcal{D}(\mathcal{J}_k) = \Pi$,

$$\mathcal{J}_k \;=\; e^{|x|^2/4}(\mathcal{H}_k - (\gamma + N/2))\, e^{-|x|^2/4}.$$

In particular, $\mathcal{J}_k$ *is symmetric in* $L^2(\mathbb{R}^N, m_k)$.

Proof. From (3.4) it is easily verified by induction that

$$[\Delta_k, E^n] = 2nE^{n-1}H + 2n(n-1)E^{n-1} \quad \text{for all} \quad n \in \mathbb{N},$$

and therefore $[\Delta_k, e^{-E/2}] = -e^{-E/2}H + \frac{1}{2}Ee^{-E/2}$. Thus on Π,

$$\begin{aligned} \mathcal{H}_k\, e^{-E/2} \;&=\; -\Delta_k e^{-E/2} + \frac{1}{2}Ee^{-E/2} \;=\; -e^{-E/2}\Delta_k + e^{-E/2}H \\ &=\; e^{-E/2}(\mathcal{J}_k + \gamma + N/2). \end{aligned}$$

Theorem 3.1. *The spaces* $L^2(\mathbb{R}^N, m_k)$ *and* $L^2(\mathbb{R}^N, w_k)$ *admit orthogonal Hilbert space decompositions into eigenspaces of the operators* $\mathcal{J}_k$ *and* $\mathcal{H}_k$ *respectively. More precisely, define*

$$V_n := \{e^{-\Delta_k/2}\,p : p \in \mathcal{P}_n\} \subset \Pi, \quad W_n := \{e^{-|x|^2/4}\,q(x), \;\; q \in V_n\} \subset \mathscr{S}(\mathbb{R}^N).$$

Then V_n *is the eigenspace of* $\mathcal{J}_k$ *corresponding to the eigenvalue* n, W_n *is the eigenspace of* $\mathcal{H}_k$ *corresponding to the eigenvalue* $n + \gamma + N/2$, *and*

$$L^2(\mathbb{R}^N, m_k) \;=\; \bigoplus_{n\in\mathbb{Z}_+} V_n\,, \quad L^2(\mathbb{R}^N, w_k) \;=\; \bigoplus_{n\in\mathbb{Z}_+} W_n\,.$$

Remark 3.1. A densely defined linear operator $(A, \mathcal{D}(A))$ in a Hilbert space H is called essentially self-adjoint, if it satisfies

(i) A is symmetric, i.e. $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x \in \mathcal{D}(A)$;
(ii) The closure $\overline{A}$ of A is selfadjoint.

In fact, every symmetric operator A in H has a unique closure $\overline{A}$ (because $A \subseteq A^*$, and the adjoint A^* is closed). If H has a countable orthonormal basis $\{v_n, n \in \mathbb{Z}_+\} \subset \mathcal{D}(A)$ consisting of eigenvectors of A corresponding to eigenvalues $\lambda_n \in \mathbb{R}$, then it is straightforward that A is essentially self-adjoint, and that the spectrum of the self-adjoint operator $\overline{A}$ is given by $\sigma(\overline{A}) = \{\lambda_n\,, n \in \mathbb{Z}_+\}$. (See for instance Lemma 1.2.2 of [8]).

In our situation, the operator $\mathcal{H}_k$ is densely defined and symmetric in $L^2(\mathbb{R}^N, w_k)$ (the first order Dunkl operators being anti-symmetric), and the same holds for $\mathcal{J}_k$ in $L^2(\mathbb{R}^N, m_k)$. The above theorem implies that $\mathcal{H}_k$ and $\mathcal{J}_k$ are essentially self-adjoint and that

$$\sigma(\overline{\mathcal{H}_k}) = \{n + \gamma + N/2, n \in \mathbb{Z}_+\}, \quad \sigma(\overline{\mathcal{J}_k}) = \mathbb{Z}_+\,.$$

Proof (of Theorem 3.1). Equation (3.4) and induction yield the commuting relations $[\rho, \Delta_k^n] = -2n\Delta_k^n$ for all $n \in \mathbb{Z}_+$, and hence

$$[\rho, e^{-\Delta_k/2}] \,=\, \Delta_k\, e^{-\Delta_k/2}.$$

If $q \in \Pi$ is arbitrary and $p := e^{\Delta_k/2}q$, it follows that

$$\rho(q) \,=\, (\rho e^{-\Delta_k/2})(p) \,=\, e^{-\Delta_k/2}\rho(p) + \Delta_k e^{-\Delta_k/2}\,p \,=\, e^{-\Delta_k/2}\rho(p) + \Delta_k q.$$

Hence for $a \in \mathbb{C}$ there are equivalent:

$$(-\Delta_k + \rho)(q) = aq \iff \rho(p) = ap \iff a = n \in \mathbb{Z}_+ \text{ and } p \in \mathcal{P}_n.$$

Thus each function from V_n is an eigenfunction of $\mathcal{J}_k$ corresponding to the eigenvalue n, and $V_n \perp V_m$ for $n \neq m$ by the symmetry of $\mathcal{J}_k$. This proves the statements for $\mathcal{J}_k$ because $\Pi = \bigoplus V_n$ is dense in $L^2(\mathbb{R}^N, m_k)$. The statements for $\mathcal{H}_k$ are then immediate by the Lemma 3.1.

3.3 Generalized Hermite polynomials

The eigenvalues of the CMS Hamiltonians $\mathcal{H}_k$ and $\mathcal{J}_k$ are highly degenerate if $N > 1$. In this section, we construct natural orthogonal bases for them. They are made up by generalizations of the classical N-variable Hermite polynomials and Hermite functions to the Dunkl setting. We follow [50], but change our normalization by a factor 2.

The starting point for our construction is the Macdonald-type identity: if $p, q \in \Pi$, then

$$[p, q]_k \,=\, \int_{\mathbb{R}^N} e^{-\Delta_k/2}p(x)e^{-\Delta_k/2}q(x)dm_k(x), \tag{3.6}$$

with the probability measure m_k defined according to (3.3). Notice that $[\,.\,,\,.\,]_k$ is a scalar product on the $\mathbb{R}$- vector space $\Pi_{\mathbb{R}}$ of polynomials with real coefficients. Let $\{\varphi_\nu\,, \nu \in \mathbb{Z}_+^N\}$ be an orthonormal basis of $\Pi_{\mathbb{R}}$ with respect to the scalar product $[.\,,.]_k$ such that $\varphi_\nu \in \mathcal{P}_{|\nu|}$. As homogeneous polynomials of different degrees are orthogonal, the φ_ν with fixed $|\nu| = n$ can for example be constructed by Gram-Schmidt orthogonalization within $\mathcal{P}_n \cap \Pi_{\mathbb{R}}$ from an arbitrary ordered real-coefficient basis. If $k = 0$, the canonical choice of the basis $\{\varphi_\nu\}$ is just $\varphi_\nu(x) := (\nu!)^{-1/2} x^\nu$.

Definition 3.1. *The generalized Hermite polynomials , $\{H_\nu\,,\ \nu \in \mathbb{Z}_+^N\}$ associated with the basis $\{\varphi_\nu\}$ on $\mathbb{R}^N$ are given by*

$$H_\nu(x) := e^{-\Delta_k/2}\varphi_\nu(x). \tag{3.7}$$

Moreover, we define the generalized Hermite functions , on $\mathbb{R}^N$ by

$$h_\nu(x) := e^{-|x|^2/4} H_\nu(x), \quad \nu \in \mathbb{Z}_+^N. \tag{3.8}$$

H_ν is a polynomial of degree $|\nu|$ satisfying $H_\nu(-x) = (-1)^{|\nu|} H_\nu(x)$ for all $x \in \mathbb{R}^N$. By virtue of (3.6), the H_ν, $\nu \in \mathbb{Z}_+^N$ form an orthonormal basis of $L^2(\mathbb{R}^N, m_k)$.

Examples 3.1. (1) *Classical multivariable Hermite polynomials.* Let $k = 0$, and choose the standard orthonormal system $\varphi_\nu(x) = (\nu!)^{-1/2} x^\nu$, with respect to $[.\,,.]_0$. The associated Hermite polynomials are given by

$$H_\nu(x) \;=\; \frac{1}{\sqrt{\nu!}} \prod_{i=1}^N e^{-\partial_i^2/2}(x_i^{\nu_i}) \;=\; \frac{2^{-|\nu|/2}}{\sqrt{\nu!}} \prod_{i=1}^N \widehat{H}_{\nu_i}(x_i/\sqrt{2}), \tag{3.9}$$

where the $\widehat{H}_n$, $n \in \mathbb{Z}_+$ are the classical Hermite polynomials on $\mathbb{R}$ defined by

$$\widehat{H}_n(x) \;=\; (-1)^n\, e^{x^2}\, \frac{d^n}{dx^n} e^{-x^2}\,.$$

(2) *The one-dimensional case.* Up to sign changes, there exists only one orthonormal basis with respect to $[\,.\,,\,.\,]_k$. The associated Hermite polynomials are given, up to multiplicative constants, by the generalized Hermite polynomials $H_n^k(x/\sqrt{2})$ on $\mathbb{R}$. These polynomials can be found in [7] and were further studied in [56] in connection with a Bose-like oscillator calculus. The H_n^k are orthogonal with respect to $|x|^{2k} e^{-|x|^2}$ and can be written as

$$\begin{cases} H_{2n}^k(x) = (-1)^n 2^{2n} n!\, L_n^{k-1/2}(x^2), \\ H_{2n+1}^k(x) = (-1)^n 2^{2n+1} n!\, x L_n^{k+1/2}(x^2); \end{cases}$$

here the L_n^α are the usual Laguerre polynomials of index $\alpha \geq -1/2$, given by

$$L_n^\alpha(x) = \frac{1}{n!} x^{-\alpha} e^x\, \frac{d^n}{dx^n}\big(x^{n+\alpha} e^{-x}\big).$$

(3) *The* A_{N-1}*-case.* There exists a natural orthogonal system $\{\varphi_\nu\}$, made up by the so-called *non-symmetric Jack polynomials.* For a multiplicity parameter $k > 0$, the associated non-symmetric Jack polynomials E_ν, $\nu \in \mathbb{Z}_+^N$, as introduced in [45] (see also [35]), are uniquely defined by the following conditions:
(i) $E_\nu(x) = x^\nu + \sum_{\mu <_P \nu} c_{\nu,\mu} x^\mu$ with $c_{\nu,\mu} \in \mathbb{R}$;
(ii) For all $\mu <_P \nu$, $(E_\nu(x), x^\mu)_k = 0$.
Here $<_P$ is a dominance order defined within multi-indices of equal total length (see [45]), and the inner product $(.,.)_k$ on $\Pi \cap \Pi_{\mathbb{R}}$ is given by

$$(f,g)_k := \int_{\mathbb{T}^N} f(z) g(\overline{z}) \prod_{i<j} |z_i - z_j|^{2k} dz$$

with $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ and dz being the Haar measure on $\mathbb{T}^N$. If f and g have different total degrees, then $(f,g)_k = 0$. The set $\{E_\nu,\, |\nu| = n\}$ forms a vector space basis of $\mathcal{P}_n \cap \Pi_{\mathbb{R}}$. It can be shown (by use of A_{N-1}-type Cherednik operators) that the Jack polynomials E_ν are also orthogonal with respect to the Dunkl pairing $[.\,,.]_k$; for details see [50]. The corresponding generalized Hermite polynomials and their symmetric counterparts have been studied in [37], [38] and in [2] - [4].

As an immediate consequence of Theorem 3.1 we obtain analogues of the classical second order differential equations for generalized Hermite polynomials and Hermite functions:

Corollary 3.1. (i) $(-\Delta_k + \sum_{i=1}^N x_i \partial_i) H_\nu = |\nu| H_\nu$.

(ii) $(-\Delta_k + \frac{1}{4}|x|^2) h_\nu = (|\nu| + \gamma + N/2) h_\nu$.

Various further useful properties of the classical Hermite polynomials and Hermite functions have extensions to our general setting. We conclude this section with a list of them. The proofs can be found in [50]. For further results on generalized Hermite polynomials, one can also see for instance [9].

Theorem 3.2. *Let* $\{H_\nu\}$ *be the Hermite polynomials associated with the basis* $\{\varphi_\nu\}$ *on* $\mathbb{R}^N$ *and let* $x, y \in \mathbb{R}^N$. *Then*

(1) *(Rodrigues formula)* $H_\nu(x) = (-1)^{|\nu|} e^{|x|^2/2} \varphi_\nu(T) e^{-|x|^2/2}$

(2) *(Generating relation)* $e^{-|y|^2/2} E_k(x,y) = \sum_{\nu \in \mathbb{Z}_+^N} H_\nu(x) \varphi_\nu(y)$

(3) *(Mehler formula) For all* $0 < r < 1$,

$$\sum_{\nu \in \mathbb{Z}_+^N} H_\nu(x) H_\nu(y) r^{|\nu|} =$$

$$\frac{1}{(1-r^2)^{\gamma+N/2}} \exp\left\{-\frac{r^2(|x|^2 + |y|^2)}{2(1-r^2)}\right\} E_k\Big(\frac{rx}{1-r^2}, y\Big).$$

The sums are absolutely convergent in both cases.

The Dunkl kernel E_k in (2) and (3) replaces the usual exponential function. It comes in via the following relation with the (arbitrary!) basis $\{\varphi_\nu\}$:

$$E_k(x,y) \,=\, \sum_{\nu\in\mathbb{Z}_+^N} \varphi_\nu(x)\varphi_\nu(y) \quad (x,y\in\mathbb{R}^N).$$

Proposition 3.1. *The generalized Hermite functions* $\{h_\nu\,,\,\nu\in\mathbb{Z}_+^N\}$ *are a basis of eigenfunctions of the Dunkl transform on* $L^2(\mathbb{R}^N,w_k)$ *with*

$$h_\nu^{\wedge k} \,=\, (-i)^{|\nu|}h_\nu.$$

4 Positivity results

4.1 Positivity of Dunkl's intertwining operator

In this section it is always assumed that $k\geq 0$. The reference is [51].

We shall say that a linear operator A on Π is *positive*, if A leaves the positive cone

$$\Pi_+ := \{p\in\Pi : p(x)\geq 0 \quad \text{for all } \ x\in\mathbb{R}^N\}$$

invariant. The following theorem is the central result of this section:

Theorem 4.1. V_k *is positive on* Π.

Once this is known, more detailed information about V_k can be obtained by its extension to the algebras A_r, which were introduced in Definition 2.5. This leads to

Theorem 4.2. *For each* $x\in\mathbb{R}^N$ *there exists a unique probability measure* μ_x^k *on the Borel-*σ*-algebra of* $\mathbb{R}^N$ *such that*

$$V_kf(x) \,=\, \int_{\mathbb{R}^N} f(\xi)\,d\mu_x^k(\xi) \qquad \textit{for all } f\in A_{|x|}. \tag{4.1}$$

The representing measures μ_x^k *are compactly supported with* $\operatorname{supp}\mu_x^k\subseteq \operatorname{co}\{gx,\,g\in G\}$, *the convex hull of the orbit of* x *under* G. *Moreover, they satisfy*

$$\mu_{rx}^k(B) \,=\, \mu_x^k(r^{-1}B), \quad \mu_{gx}^k(B) \,=\, \mu_x^k(g^{-1}(B)) \tag{4.2}$$

for each $r>0,\ g\in G$ *and each Borel set* $B\subseteq\mathbb{R}^N$.

The proof of Theorem 4.1 affords several steps, the crucial one being a reduction from the N-dimensional to a one-dimensional problem. We shall give an outline, but beforehand we turn to the proof of Theorem 4.2.

Proof (of Theorem 4.2). Fix $x \in \mathbb{R}^N$ and put $r = |x|$. Then the mapping

$$\Phi_x : f \mapsto V_k f(x)$$

is a bounded linear functional on A_r, and Theorem 4.1 implies that it is positive on the dense subalgebra Π of A_r, i.e. $\Phi_x(|p|^2) \geq 0$. Consequently, Φ_x is a positive functional on the full Banach-$*$-algebra A_r. There exists a representation theorem of Bochner for positive functionals on commutative Banach-$*$-algebras (see for instance Theorem 21.2 of [21]). It implies in our case that there exists a unique measure $\nu_x \in M_b^+(\Delta_S(A_r))$ such that

$$\Phi_x(f) = \int_{\Delta_S(A_r)} \widehat{f}(\varphi)\, d\nu_x(\varphi) \qquad \text{for all } f \in A_r,$$

with $\widehat{f}$ the Gelfand transform of f. Keeping Exercise 7 in mind, one obtains representing measures μ_x^k supported in the ball B_r; the sharper statement on the support is obtained by results of [30]. The remaining statements are easy.

The key for the proof of Theorem 4.1 is a characterization of positive semigroups on polynomials which are generated by degree-lowering operators. We call a linear operator A on Π *degree-lowering*, if $deg(Ap) < deg(p)$ for all Π. Again, the exponential $e^A \in \mathrm{End}(\Pi^N)$ is defined by a terminating power-series, and it can be considered as a linear operator on each of the finite dimensional spaces $\{p \in \Pi : deg(p) \leq m\}$. Important examples of degree-lowering operators are linear operators which are homogeneous of some degree $-n < 0$, such as Dunkl operators. The following key result characterizes positive semigroups generated by degree-lowering operators; it is an adaption of a well-known Hille-Yosida type characterization theorem for so called Feller-Markov semigroups which will be discussed a little later in our course, see Theorem 4.7.

Theorem 4.3. *Let A be a degree-lowering linear operator on Π. Then the following statements are equivalent:*
(1) *e^{tA} is positive on Π for all $t \geq 0$.*
(2) *A satisfies the "positive minimum principle"*

(M) *For every $p \in \Pi_+$ and $x_0 \in \mathbb{R}^N$, $p(x_0) = 0$ implies $Ap\,(x_0) \geq 0$.*

Exercise 9.

(1) Prove implication (1) $\Rightarrow$ (2) of this theorem.
(2) Verify that the (usual) Laplacian Δ satisfies the positive minimum principle (M). Can you extend this result to the Dunkl Laplacian Δ_k? (C.f. Exercise 12!)

Let us now outline the proof of Theorem 4.1. We consider the generalized Laplacian Δ_k associated with G and k, which is homogeneous of degree -2 on Π. With the notation introduced in (2.2), it can be written as

$$\Delta_k = \Delta + L_k \quad \text{with } L_k = 2\sum_{\alpha\in R_+} k(\alpha)\delta_\alpha\,. \tag{4.3}$$

Here δ_α acts in direction α only.

Theorem 4.4. *The operator* $e^{-\Delta/2}e^{\Delta_k/2}$ *is positive on* Π.

Proof. We shall deduce this statement from a positivity result for a suitable semigroup. For this, we employ Trotter's product formula, which works for degree-lowering operators just as on finite-dimensional vector spaces: If A, B are degree-lowering linear operators on Π, then

$$e^{A+B}p\,(x) = \lim_{n\to\infty}(e^{A/n}e^{B/n})^n p\,(x).$$

Thus, we can write

$$\begin{aligned} e^{-\Delta/2}e^{\Delta_k/2}\,p\,(x) &= e^{-\Delta/2}e^{\Delta/2+L_k/2}\,p\,(x) = \lim_{n\to\infty} e^{-\Delta/2}\big(e^{\Delta/2n}\,e^{L_k/2n}\big)^n p\,(x) \\ &= \lim_{n\to\infty}\prod_{j=1}^n \big(e^{-(1-j/n)\cdot\Delta/2}\,e^{L_k/2n}\,e^{(1-j/n)\cdot\Delta/2}\big)p\,(x). \end{aligned}$$

It therefore suffices to verify that the operators

$$e^{-s\Delta}e^{tL_k}e^{s\Delta} \quad (s,t\geq 0)$$

are positive on Π. Consider s fixed, then

$$e^{-s\Delta}e^{tL_k}e^{s\Delta} = e^{tA} \quad \text{with } A = e^{-s\Delta}L_k e^{s\Delta}\,.$$

It is easily checked that A is degree-lowering. Hence, in view of Theorem 4.3, it remains to show that A satisfies the positive minimum principle (M). We may write

$$A = e^{-s\Delta}L_k e^{s\Delta} = 2\sum_{\alpha\in R_+} k(\alpha)e^{-s\partial_\alpha^2}\delta_\alpha\, e^{s\partial_\alpha^2};$$

here it was used that δ_α acts in direction α only. It can now be checked by direct computation that the one-dimensional operators $e^{-s\partial_\alpha^2}\delta_\alpha e^{s\partial_\alpha^2}$ satisfy (M), and as the $k(\alpha)$ are non-negative, this must be true for A as well.

Proof (of Theorem 4.1). Notice first that

$$[V_k\, p, q]_k = [p,q]_0 \quad \text{for all } p,q\in\Pi. \tag{4.4}$$

In fact, for $p,q\in\mathcal{P}_n$ with $n\in\mathbb{Z}_+$, one obtains

$$[V_k\, p, q]_k = [q, V_k\, p]_k = q(T)(V_k\, p) = V_k(q(\partial)p) = q(\partial)(p) = [p,q]_0\,;$$

here the characterizing properties of V_k and the fact that $q(\partial)(p)$ is a constant have been used. For general $p,q\in\Pi$, (4.4) then follows from the orthogonality of the spaces $\mathcal{P}_n$ with respect to both pairings.

Combining the Macdonald-type identity (2.5) with part (4.4), we obtain for all $p, q \in \Pi$ the identity

$$c_k^{-1}\int_{\mathbb{R}^N} e^{-\Delta_k/2}(V_kp)e^{-\Delta_k/2}q\, e^{-|x|^2/2}\, w_k(x)dx \,=\, c_0^{-1}\int_{\mathbb{R}^N} e^{-\Delta/2}p\, e^{-\Delta/2}q\, e^{-|x|^2/2}dx.$$

As $e^{-\Delta_k/2}(V_k\, p) = V_k(e^{-\Delta/2}p)$, and as we may also replace p by $e^{\Delta/2}p$ and q by $e^{\Delta_k/2}q$ in the above identity, it follows that for all $p, q \in \Pi$

$$c_k^{-1}\int_{\mathbb{R}^N} V_kp\, q\, e^{-|x|^2/2}\, w_k(x)dx \,=\, c_0^{-1}\int_{\mathbb{R}^N} p\, e^{-\Delta/2}e^{\Delta_k/2}q\, e^{-|x|^2/2}dx. \tag{4.5}$$

Due to Theorem 4.4, the right side of (4.5) is non-negative for all $p, q \in \Pi_+$. From this, the assertion can be deduced by standard density arguments (Π is dense in $L^2(\mathbb{R}^N, e^{-|x|^2/4}w_k(x)dx)$).

Corollary 4.1. *For each $y \in \mathbb{C}^N$, the function $x \mapsto E_k(x,y)$ has the Bochner-type representation*

$$E_k(x,y) \,=\, \int_{\mathbb{R}^N} e^{\langle \xi,\, y\rangle} d\mu_x^k(\xi), \tag{4.6}$$

where the μ_x^k are the representing measures from Theorem 4.2. In particular, E_k satisfies the estimates stated in Proposition 2.6, and

$$E_k(x,y) > 0 \quad \textit{for all } x, y \in \mathbb{R}^N.$$

Analogous statements hold for the k-Bessel function J_k.

In those cases where the generalized Bessel functions $J_k(.,y)$ allow an interpretation as the spherical functions of a Cartan motion group, the Bochner representation of these functions is an immediate consequence of Harish-Chandra's theory ([28]). There are, however, no group-theoretical interpretations known for the kernel E_k so far.

4.2 Heat kernels and heat semigroups

We start with a motivation: Consider the following initial-value problem for the classical heat equation in $\mathbb{R}^N$:

$$\begin{cases} \Delta u - \partial_t u \,= 0 \quad \text{on } \mathbb{R}^N \times (0,\infty), \\ u(\,.\,,0) = f \end{cases} \tag{4.7}$$

with initial data $f \in C_0(\mathbb{R}^N)$, the space of continuous functions on $\mathbb{R}^N$ which vanish at infinity. (We could equally take data from $C_b(\mathbb{R}^N)$, but $C_0(\mathbb{R}^N)$ is

more convenient in the following considerations). The basic idea to solve (4.7) is to carry out a Fourier transform with respect to x. This yields the candidate

$$u(x,t) = g_t * f(x) = \int_{\mathbb{R}^N} g_t(x-y)f(y)dy \quad (t>0), \tag{4.8}$$

where g_t is the Gaussian kernel

$$g_t(x) = \frac{1}{(4\pi t)^{N/2}}\, e^{-|x|^2/4t}.$$

It is a well-known fact from classical analysis that (4.8) is in fact the unique bounded solution within the class $C^2(\mathbb{R}^N \times (0,\infty)) \cap C(\mathbb{R}^N \times [0,\infty))$.

Exercise 10. Show that $H(t)f(x) := g_t * f(x)$ for $t>0$, $H(0) := id$ defines a strongly continuous contraction semigroup on the Banach space $(C_0(\mathbb{R}^N), \|.\|_\infty)$ in the sense of the definition given below.
Hint: Once contractivity is shown, it suffices to check the continuity for functions from the Schwartz space $\mathscr{S}(\mathbb{R}^N)$. For this, use the Fourier inversion theorem.

Definition 4.1. *Let X be a Banach space. A one-parameter family $(T(t))_{t\geq 0}$ of bounded linear operators on X is called a strongly continuous semigroup on X, if it satisfies*

(i) $T(0) = id_X$, $\quad T(t+s) = T(t)T(s)$ *for all* $t,s \geq 0$
(ii) *The mapping $t \mapsto T(t)x$ is continuous on $[0,\infty)$ for all $x \in X$.*

A strongly continuous semigroup is called a contraction semigroup, if $\|T(t)\| \leq 1$ for all $t \geq 0$.

Let $L(X)$ denote the space of bounded linear operators in X. If $A \in L(X)$, then

$$e^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n \in L(X)$$

defines a strongly continuous semigroup on X (this one is even continuous with respect to the uniform topology on $L(X)$). We obviously have

$$A = \lim_{t\downarrow 0} \frac{1}{t}(e^{tA} - id) \quad \text{in } L(X).$$

Definition 4.2. *The generator of a strongly continuous semigroup $(T(t))_{t\geq 0}$ in X is defined by*

$$\begin{aligned} Ax &:= \lim_{t\downarrow 0} \frac{1}{t}(T(t)x - x), \quad \textit{with domain} \\ \mathcal{D}(A) &:= \{x \in X : \lim_{t\downarrow 0} \frac{1}{t}(T(t)x - x) \textit{ exists in } X\}. \end{aligned} \tag{4.9}$$

Theorem 4.5. *The generator A of $(T(t))_{t\geq 0}$ is densely defined and closed.*

An important issue in the theory of operator semigroups and evolution equations are criteria which characterize generators of strongly continuous semigroups.

Let us return to the Dunkl setting. As before, Δ_k denotes the Dunkl Laplacian associated with a finite reflection group on $\mathbb{R}^N$ and some multiplicity function $k \geq 0$, and the index γ is defined according to (2.4). We are going to consider the following initial-value problem for the Dunkl-type heat operator $\Delta_k - \partial_t$:

Find $u \in C^2(\mathbb{R}^N \times (0,\infty))$ which is continuous on $\mathbb{R}^N \times [0,\infty)$ and satisfies

$$\begin{cases} (\Delta_k - \partial_t)\, u \,=\, 0 & \text{on } \mathbb{R}^N \times (0,\infty), \\ u(\,.\,,0) \,=\, f & \in C_b(\mathbb{R}^N). \end{cases} \tag{4.10}$$

The solution of this problem is given, just as in the classical case $k = 0$, in terms of a positivity-preserving semigroup. We shall essentially follow the treatment of [50].

Lemma 4.1. *The function*

$$F_k(x,t) := \frac{1}{(2t)^{\gamma+N/2} c_k} e^{-|x|^2/4t}$$

solves the generalized heat equation $\Delta_k u - \partial_t u = 0$ on $\mathbb{R}^N \times (0,\infty)$.

Proof. A short calculation. Use the product rule (2.1) as well as the identity $\sum_{i=1}^N T_i(x_i) = N + 2\gamma$.

F_k generalizes the fundamental solution for the classical heat equation which is given by $F_0(x,t) = g_t(x)$ (as defined above). It is easily checked that

$$\int_{\mathbb{R}^N} F_k(x,t)\, w_k(x) dx \,=\, 1 \quad \text{for all } t > 0.$$

In order to solve (4.10), it suggests itself to apply the Dunkl transform under suitable decay assumptions on the initial data. In the classical case, the heat kernel $g_t(x-y)$ on $\mathbb{R}^N$ is obtained from the fundamental solution simply by translations. In the Dunkl setting, it is still possible to define a generalized translation which matches the action of the Dunkl transform, i.e. makes it a homomorphism on suitable function spaces.

The notion of a *generalized translation* in the Schwartz space $\mathscr{S}(\mathbb{R}^N)$ is as follows (c.f. [50]):

$$\tau_y f(x) := \frac{1}{c_k} \int_{\mathbb{R}^N} \widehat{f}^{\,k}(\xi)\, E_k(ix,\xi) E_k(iy,\xi)\, w_k(\xi) d\xi; \quad y \in \mathbb{R}^N. \tag{4.11}$$

In the same way, this could be done in $L^2(\mathbb{R}^N, w_k)$. A powerful extension to $C^\infty(\mathbb{R}^N)$ is due to Trimèche [59]. Note that in case $k = 0$, we simply

have $\tau_y f(x) = f(x+y)$. In the rank-one case, the above translation coincides with the convolution on a so-called signed hypergroup structure which was defined in [49]; see also [56]. Similar structures are not yet known in higher rank cases. Clearly, $\tau_y f(x) = \tau_x f(y)$; moreover, the inversion theorem for the Dunkl transform assures that $\tau_0 f = f$ and

$$(\tau_y f)^{\wedge k}(\xi) = E_k(iy,\xi)\widehat{f}^k(\xi).$$

From this it is not hard to see that $\tau_y f$ belongs to $\mathscr{S}(\mathbb{R}^N)$ again. Let us now consider the "fundamental solution" $F_k(.\,,t)$ for $t > 0$. A short calculation, using the reproducing property Proposition 2.7(2), shows that

$$\widehat{F}_k^k(\xi,t) \,=\, c_k^{-1} e^{-t|\xi|^2}. \tag{4.12}$$

By the quoted reproducing formula one therefore obtains from (4.11) the representation

$$\tau_{-y}F_k(x,t) \,=\, \frac{1}{(2t)^{\gamma+N/2}c_k} e^{-(|x|^2+|y|^2)/4t}\, E_k\Big(\frac{x}{\sqrt{2t}},\frac{y}{\sqrt{2t}}\Big).$$

This motivates the following

Definition 4.3. *The generalized heat kernel* Γ_k *is defined by*

$$\Gamma_k(t,x,y) := \frac{1}{(2t)^{\gamma+N/2}c_k}\, e^{-(|x|^2+|y|^2)/4t}\, E_k\Big(\frac{x}{\sqrt{2t}},\frac{y}{\sqrt{2t}}\Big), \quad x,y\in\mathbb{R}^N,\ t>0.$$

Notice in particular that $\Gamma_k > 0$ (thanks to Corollary 4.1) and that $y \mapsto \Gamma_k(t,x,y)$ belongs to $\mathscr{S}(\mathbb{R}^N)$ for fixed x and t. We collect a series of further fundamental properties of this kernel which are all more or less straightforward.

Lemma 4.2. *The heat kernel* Γ_k *has the following properties:*

(1) $\Gamma_k(t,x,y) \,=\, c_k^{-2}\displaystyle\int_{\mathbb{R}^N} e^{-t|\xi|^2}\, E_k(ix,\xi)\, E_k(-iy,\xi)\, w_k(\xi)d\xi\,.$

(2) $\displaystyle\int_{\mathbb{R}^N} \Gamma_k(t,x,y)\, w_k(y)dy \,=\, 1.$

(3) $\Gamma_k(t,x,y) \,\leq\, \dfrac{1}{(2t)^{\gamma+N/2}c_k}\, \max_{g\in G} e^{-|gx-y|^2/4t}\,.$

(4) $\Gamma_k(t+s,x,y) \,=\, \displaystyle\int_{\mathbb{R}^N} \Gamma_k(t,x,z)\,\Gamma_k(s,y,z)\, w_k(z)dz.$

(5) *For fixed* $y\in\mathbb{R}^N$*, the function* $u(x,t) := \Gamma_k(t,x,y)$ *solves the generalized heat equation* $\Delta_k u = \partial_t u$ *on* $\mathbb{R}^N \times (0,\infty)$.

Proof. (1) is clear from the definition of generalized translations. For details concerning (2) see [50]. (3) follows from our estimates on E_k, while (4) is obtained by inserting (1) for one of the kernels in the integral. Finally, (5) is obtained from differentiating (1) under the integral. For details see again [50].

Definition 4.4. *For* $f \in C_b(\mathbb{R}^N)$ *and* $t \geq 0$ *set*

$$H(t)f(x) := \begin{cases} \displaystyle\int_{\mathbb{R}^N} \Gamma_k(t,x,y)f(y)\, w_k(y)dy & \text{if } t > 0, \\ f(x) & \text{if } t = 0 \end{cases} \tag{4.13}$$

Notice that the decay of Γ_k assures the convergence of the integral. The properties of the operators $H(t)$ are most easily described on the Schwartz space $\mathscr{S}(\mathbb{R}^N)$. The following theorem is completely analogous to the classical case.

Theorem 4.6. *Let* $f \in \mathscr{S}(\mathbb{R}^N)$. *Then* $u(x,t) := H(t)f(x)$ *solves the initial-value problem* (4.10). *Moreover,* $H(t)f$ *has the following properties:*
(1) $H(t)f \in \mathscr{S}(\mathbb{R}^N)$ *for all* $t > 0$.
(2) $H(t+s)\,f \,=\, H(t)H(s)f$ *for all* $s, t \geq 0$.
(3) $\|H(t)f - f\|_{\infty,\mathbb{R}^N} \to 0$ *as* $t \to 0$.

Proof. (Sketch) By use of Lemma 4.2 (1) and Fubini's theorem, we write

$$u(x,t) \,=\, H(t)f(x) \,=\, c_k^{-1} \int_{\mathbb{R}^N} e^{-t|\xi|^2} \widehat{f}^{\,k}(\xi)\, E_k(ix,\xi)\, w_k(\xi)d\xi \quad (t > 0). \tag{4.14}$$

In view of the inversion theorem for the Dunkl transform, this holds for $t = 0$ as well. Properties (1) and (3) as well as the differential equation are now easy consequences. Part (2) follows from the reproducing formula for Γ_k (Lemma 4.2 (4)).

Exercise 11. Carry out the details in the proof of Theorem 4.6.

We know that the heat kernel Γ_k is positive; this implies that $H(t)f \geq 0$ if $f \geq 0$.

Definition 4.5. *Let* Ω *be a locally compact Hausdorff space. A strongly continuous semigroup* $(T(t))_{t\geq 0}$ *on* $(C_0(\Omega), \|.\|_\infty)$ *is called a Feller-Markov semigroup, if it is contractive and positive, i.e.* $f \geq 0$ *on* Ω *implies that* $T(t)f \geq 0$ *on* Ω *for all* $t \geq 0$.

We shall prove that the linear operators $H(t)$ on $\mathscr{S}(\mathbb{R}^N)$ extend to a Feller-Markov semigroup on the Banach space $(C_0(\mathbb{R}^N), \|.\|_\infty)$. This could be done by direct calculations similar to the usual procedure for the classical heat semigroup, relying on the positivity of the kernel Γ_k. We do however prefer to give a proof which does not require this rather deep result, but works on the level of the tentative generator. The tool is the following useful variant of the Lumer-Phillips theorem, which characterizes Feller-Markov semigroups in terms of a "positive maximum principle", see e.g. [34], Theorem 17.11. In fact, this Theorem motivated the positive minimum principle of Theorem 4.3 in the positivity-proof for V_k.

Theorem 4.7. *Let $(A, \mathcal{D}(A))$ be a densely defined linear operator in $(C_0(\Omega), \|.\|_\infty)$. Then A is closable, and its closure $\overline{A}$ generates a Feller-Markov semigroup on $C_0(\Omega)$, if and only if the following conditions are satisfied:*

(i) *If $f \in \mathcal{D}(A)$ then also $\overline{f} \in \mathcal{D}(A)$ and $A(\overline{f}) = \overline{A(f)}$.*
(ii) *The range of $\lambda id - A$ is dense in $C_0(\Omega)$ for some $\lambda > 0$.*
(iii) *If $f \in \mathcal{D}(A)$ is real-valued with a non-negative maximum in $x_0 \in \Omega$, i.e. $0 \leq f(x_0) = \max_{x\in\Omega} f(x)$, then $Af(x_0) \leq 0$. (Positive maximum principle).*

We consider the Dunkl Laplacian Δ_k as a densely defined linear operator in $C_0(\mathbb{R}^N)$ with domain $\mathscr{S}(\mathbb{R}^N)$. The following Lemma implies that it satisfies the positive maximum principle:

Lemma 4.3. *Let $\Omega \subseteq \mathbb{R}^N$ be open and G-invariant. If a real-valued function $f \in C^2(\Omega)$ attains an absolute maximum at $x_0 \in \Omega$, i.e. $f(x_0) = \sup_{x\in\Omega} f(x)$, then*

$$\Delta_k f(x_0) \leq 0.$$

Exercise 12. Prove this lemma in the case that $\langle \alpha, x_0 \rangle \neq 0$ for all $\alpha \in R$. (If $\langle \alpha, x_0 \rangle = 0$ for some $\alpha \in R$, one has to argue more carefully; for details see [50].)

Theorem 4.8. *The operators $(H(t))_{t\geq 0}$ define a Feller-Markov semigroup on $C_0(\mathbb{R}^N)$. Its generator is the closure $\overline{\Delta}_k$ of $(\Delta_k, \mathscr{S}(\mathbb{R}^N))$. This semigroup is called the generalized heat semigroup on $C_0(\mathbb{R}^N)$.*

Proof. In the first step, we check that Δ_k (with domain $\mathscr{S}(\mathbb{R}^N)$) satisfies the conditions of Theorem 4.7: Condition (i) is obvious and (iii) is an immediate consequence of the previous lemma. Condition (ii) is also satisfied, because $\lambda id - \Delta_k$ maps $\mathscr{S}(\mathbb{R}^N)$ onto itself for each $\lambda > 0$; this follows from the fact that the Dunkl transform is a homeomorphism of $\mathscr{S}(\mathbb{R}^N)$ and $((\lambda I - \Delta_k)f)^{\wedge k}(\xi) = (\lambda + |\xi|^2)\widehat{f}^k(\xi)$. Theorem 4.7 now implies that Δ_k is closable, and that its closure $\overline{\Delta}_k$ generates a Feller-Markov semigroup $(T(t))_{t\geq 0}$. It remains to show that $T(t) = H(t)$ on $C_0(\mathbb{R}^N)$. Let first $f \in \mathscr{S}(\mathbb{R}^N)$. From basic facts in semigroup theory, it follows that the function $t \mapsto T(t)f$ is the unique solution of the so-called abstract Cauchy problem

$$\begin{cases} \dfrac{d}{dt}u(t) = \overline{\Delta}_k u(t) \quad \text{for } t > 0, \\ u(0) = f \end{cases} \tag{4.15}$$

within the class of all (strongly) continuously differentiable functions u on $[0, \infty)$ with values in $(C_0(\mathbb{R}^N), \|.\|_\infty)$. It is easily seen from Theorem 4.6, and in particular from formula (4.14), that $t \mapsto H(t)f$ satisfies these conditions. Hence $T(t) = H(t)$ on $\mathscr{S}(\mathbb{R}^N)$. This easily implies that $\Gamma_k \geq 0$ (which we did not presuppose for the proof!), and therefore the operators $H(t)$ are also contractive on $C_0(\mathbb{R}^N)$. A density argument now finishes the proof.

Based on this result, it is checked by standard arguments that for data $f \in C_b(\mathbb{R}^N)$, the function $u(x,t) := H(t)f(x)$ solves the initial-value problem (4.10). Uniqueness results are established by means of maximum principles, just as with the classical heat equation. Moreover, the heat semigroup $(H(t))_{t\geq 0}$ can also be defined (by means of (4.13)) on the Banach spaces $L^p(\mathbb{R}^N, w_k), 1 \leq p < \infty$. In case $p = 2$, the following is easily seen by use of the Dunkl transform:

Proposition 4.1. [50] *The operator $(\Delta_k, \mathscr{S}(\mathbb{R}^N))$ in $L^2(\mathbb{R}^N, w_k)$ is densely defined and closable. Its closure generates a strongly continuous and positivity-preserving contraction semigroup on $L^2(\mathbb{R}^N, w_k)$ which is given by*

$$H(t)f(x) = \int_{\mathbb{R}^N} \Gamma_k(t,x,y)f(y)w_k(y)dy\,, \quad (t > 0).$$

Theorem 4.8 was the starting point in [55] to construct an associated Feller-Markov process on $\mathbb{R}^N$ which can be considered a generalization of the usual Brownian motion. The transition probabilities of this process are defined in terms of a semigroup of Markov kernels of $\mathbb{R}^N$, as follows: For $x \in \mathbb{R}^N$ and a Borel set $A \in \mathscr{B}(\mathbb{R}^N)$ put

$$P_t(x,A) := \int_A \Gamma_k(t,x,y)w_k(y)dy \quad (t > 0), \quad P_0(x,A) := \delta_x(A),$$

with δ_x denoting the point measure in $x \in \mathbb{R}^N$. Then $(P_t)_{t\geq 0}$ is a semigroup of Markov kernels on $\mathbb{R}^N$ in the following sense:

(1) Each P_t is a Markov kernel, and for all $s,t \geq 0, x \in \mathbb{R}^N$ and $A \in \mathscr{B}(\mathbb{R}^N)$,

$$P_s \circ P_t(x,A) := \int_{\mathbb{R}^N} P_t(z,A)\, P_s(x,dz) \;=\; P_{s+t}(x,A).$$

(2) The mapping $[0,\infty) \to M^1(\mathbb{R}^N)$, $t \mapsto P_t(0,\,.\,)$, is continuous with respect to the $\sigma(M^1(\mathbb{R}^N, C_b(\mathbb{R}^N))$-topology.

Moreover, the semigroup $(P_t)_{t\geq 0}$ has the following particular property:

(3) $P_t(x,.\,)^{\wedge k}(\xi) = E_k(-ix,\xi)\, P_t(0,\,.\,)^{\wedge k}(\xi)$ for all $\xi \in \mathbb{R}^N$,

hereby the Dunkl transform of the probability measures $P_t(x,\,.\,)$ is defined by

$$P_t(x,\,.\,)^{\wedge k}(\xi) := \int_{\mathbb{R}^N} E_k(-i\xi,x)\, P_t(x,d\xi).$$

The proof of (1) – (3) is straightforward by the properties of Γ_k and Theorem 4.8.

In the classical case $k = 0$, property (3) is equivalent to $(P_t)_{t\geq 0}$ being translation-invariant, i.e.

$$P_t(x+y, A+y) \;=\; P_t(x,A) \quad \text{for all } y \in \mathbb{R}^N.$$

In our general setting, a positivity-preserving translation on $M^1(\mathbb{R}^N)$ cannot be expected (and does definitely not exist in the rank-one case according to [49]). Property (3) thus serves as a substitute for translation-invariance. The reader can see [55] for a study of the semigroup $(P_t)_{t\geq 0}$ and the associated Feller-Markov process.

5 Asymptotic analysis for the Dunkl kernel

This final section deals with the asymptotic behavior of the Dunkl kernel E_k with $k \geq 0$ when one of its arguments is fixed and the other tends to infinity either within a Weyl chamber of the associated reflection group, or within a suitable complex domain. These results are contained in [54]. They generalize the well-known asymptotics of the confluent hypergeometric function ${}_1F_1$ to the higher-rank setting. One motivation to study the asymptotics of E_k is to determine the asymptotic behavior of the Dunkl-type heat kernel Γ_k for short times. Partial results in this direction were obtained in [52].

Recall from Proposition 4.1 that Γ_k is the kernel of the generalized heat semigroup in the weighted $L^2(\mathbb{R}^N, w_k)$. We want to compare it with the free Gaussian kernel Γ_0. For this, it is appropriate to transfer the semigroup $(H(t))_{t\geq 0}$ from $L^2(\mathbb{R}^N, w_k)$ to the unweighted space $L^2(\mathbb{R}^N)$, which leads to the strongly continuous contraction semigroup

$$\widetilde{H}(t)f := w_k^{1/2} H(t)(w_k^{-1/2} f), \quad f \in L^2(\mathbb{R}^N).$$

The corresponding renormalized heat kernel is given by

$$\widetilde{\Gamma}_k(t,x,y) := \sqrt{w_k(x)w_k(y)}\,\Gamma_k(t,x,y).$$

The generator of $(\widetilde{H}(t))_{t\geq 0}$ is the gauge-transformed version of the Dunkl Laplacian discussed in connection with CMS-models,

$$\mathcal{F}_k \,=\, \Delta - 2\sum_{\alpha\in R_+} \frac{k(\alpha)}{\langle \alpha, x\rangle^2}(k(\alpha) - \sigma_\alpha)$$

(with suitable domain). $\mathcal{F}_k$ can be considered a perturbation of the Laplacian Δ. This suggests that within the Weyl chambers of G, the heat kernel $\widetilde{\Gamma}_k(t,x,y)$ should not “feel” the reflecting hyperplanes and behave for short times like the free Gaussian kernel $\Gamma_0(t,x,y) \,=\, g_t(x-y)$, in other words, we have the conjecture

$$\lim_{t\downarrow 0} \frac{\sqrt{w_k(x)w_k(y)}\,\Gamma_k(t,x,y)}{\Gamma_0(t,x,y)} \,=\, 1 \tag{5.1}$$

provided x and y belong to the same (open) Weyl chamber. In [52], this could be proven true only for a restricted range of arguments x, y, and by rather technical methods (completely different from those below).

Example 5.1. **The rank-one case.** Here E_k is explicitly known. According to Example 2.1,

$$E_k(z,w) \;=\; e^{zw}\cdot {}_1F_1(k, 2k+1, -2zw), \quad z,w\in\mathbb{C}.$$

The confluent hypergeometric function ${}_1F_1$ has well-known asymptotic expansions in the sectors

$$S_+ = \{z\in\mathbb{C} : -\pi/2 < \arg(z) < 3\pi/2\},$$
$$S_- = \{z\in\mathbb{C} : -3\pi/2 < \arg(z) < \pi/2\},$$

see for instance [1]. They are of the form

$$\begin{aligned}{}_1F_1(k, 2k+1, z) \;&= \frac{\Gamma(2k+1)}{\Gamma(k)} e^z z^{-k-1}(1+\mathcal{O}(\frac{1}{|z|}))\\ &\quad + \frac{\Gamma(2k+1)}{\Gamma(k+1)} e^{\pm i\pi k} z^{-k}(1+\mathcal{O}(\frac{1}{|z|})),\end{aligned}$$

with $\pm$ for $z\in S_\pm$. Specializing to the right half plane $H=\{z\in\mathbb{C}: \mathrm{Re} z\geq 0\}$ we thus obtain

$$\lim_{zw\to\infty, zw\in H} (zw)^k e^{-zw} E_k(z,w) \;=\; \frac{\Gamma(2k+1)}{2^k\Gamma(k+1)}.$$

Let us now turn to the general case of an arbitrary reflection group and multiplicity parameter $k\geq 0$. Let C denote the Weyl chamber attached with the positive subsystem R_+,

$$C = \{x\in\mathbb{R}^N : \langle\alpha, x\rangle > 0 \text{ for all } \alpha\in R_+\},$$

and for $\delta > 0$,

$$C_\delta := \{x\in C : \langle\alpha, x\rangle \;>\; \delta|x| \text{ for all } \alpha\in R_+\}.$$

The main result given here is the following asymptotic behavior, uniform for the variable tending to infinity in cones C_δ:

Theorem 5.1. *There exists a constant non-zero vector* $v=(v_g)_{g\in G}\in\mathbb{C}^{|G|}$ *such that for all* $y\in C$, $g\in G$ *and each* $\delta>0$,

$$\lim_{|x|\to\infty,\, x\in C_\delta} \sqrt{w_k(x)w_k(y)}\, e^{-i\langle x, gy\rangle} E_k(ix, gy) \;=\; v_g.$$

Notice that one variable is being fixed. A locally uniform result with respect to both variables should be true, but is open yet. Also, the explicit values of the constants v_g – apart from v_e – are not known. We come back to this point later. An immediate consequence of Theorem 5.1 is the following ray asymptotic for the Dunkl kernel (already conjectured in [15]):

Corollary 5.1. *For all* $x, y \in C$ *and* $g \in G$,

$$\lim_{t\to\infty} t^\gamma\, e^{-it\,\langle x, gy\rangle} E_k(itx, gy) \;=\; \frac{v_g}{\sqrt{w_k(x)w_k(y)}}\,,$$

the convergence being locally uniform with respect to the parameter x.

In the particular case $g = e$ (the unit of G), this latter result can be extended to a larger range of complex arguments by use of the Phragmén-Lindelöf principle for the right half plane H (see e.g. [58]):

Proposition 5.1. *Suppose* $f: H \to \mathbb{C}$ *is analytic and regular in* $H \cap \{z \in \mathbb{C} : |z| > R\}$ *for some* $R > 0$ *with* $\lim_{t\to\infty} f(it) = a$, $\lim_{t\to\infty} f(-it) = b$ *and for each* $\delta > 0$,

$$f(z) = \mathcal{O}(e^{\delta|z|}) \quad \text{as } z \to \infty \text{ within } H.$$

Then $a = b$ *and* $f(z) \to a$ *uniformly as* $z \to \infty$ *in* H.

Theorem 5.2. *Let* $x, y \in C$. *Then*

$$\lim_{z\to\infty, z\in H} z^\gamma e^{-z\langle x,y\rangle} E_k(zx, y) \;=\; \frac{i^\gamma v_e}{\sqrt{w_k(x)w_k(y)}}.$$

Here z^γ *is the holomorphic branch in* $\mathbb{C} \setminus \{x \in \mathbb{R} : x \leq 0\}$ *with* $1^\gamma = 1$.

Proof. Consider

$$G(z) := z^\gamma e^{-z\langle x,y\rangle} E_k(zx, y).$$

The estimate of Proposition 2.6 on E_k implies that G satisfies the required growth bound in Proposition 5.1 (here it is of importance that x and y lie in the same Weyl chamber), and Corollary 5.1 assures that G has limits along the boundary lines of H.

When restricted to real arguments, Theorem 5.2 implies the above stated short-time asymptotic for the heat kernel Γ_k:

Corollary 5.2. *For all* $x, y \in C$,

$$\lim_{t\downarrow 0} \frac{\sqrt{w_k(x)\, w_k(y)}\;\Gamma_k(t, x, y)}{\Gamma_0(t, x, y)} \;=\; 1\,.$$

Hereby the precise value of the limit follows from the results of [52]. Along with it, we thus obtain the value of v_e:

$$v_e \;=\; i^{-\gamma}\frac{c_k}{c_0}.$$

We shall now give an outline of the proof of Theorem 5.1. It is based on the analysis of an associated system of first order differential equations, which is derived from the eigenfunction characterization (2.7) of E_k. This approach

goes back to [30], where it was used to obtain exponential estimates for the Dunkl kernel. Put

$$\mathbb{R}^N_{reg} := \mathbb{R}^N \setminus \{\langle\alpha\rangle^\perp, \alpha \in R\}$$

and define

$$\varphi(x,y) = \sqrt{w_k(x)w_k(y)}\, e^{-i\langle x,y\rangle} E_k(ix,y), \quad x,y \in \mathbb{R}^N.$$

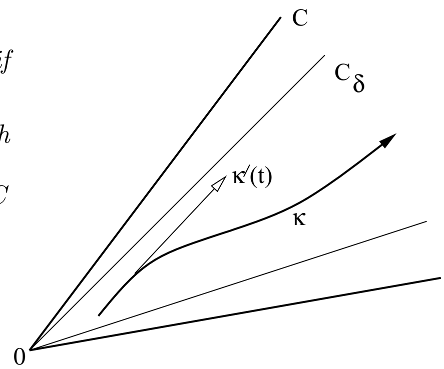
Observe that φ is symmetric in its arguments. We have to study the asymptotic behavior of $x \mapsto \varphi(x,y)$ along curves in C, with the second component $y \in \mathbb{R}^N_{reg}$ being fixed. Let us introduce the auxiliary vector field $F = (F_g)_{g\in G}$ on $\mathbb{R}^N \times \mathbb{R}^N$ by

$$F_g(x,y) := \varphi(x,gy).$$

For fixed y, we consider F along a differentiable curve $\kappa : (0,\infty) \to C$. The eigenfunction characterization of E_k then translates into a first order ordinary differential equation for $t \mapsto F(\kappa(t),y)$. Below, we shall determine the asymptotic behavior of its solutions, provided κ is admissible in the following sense:

Definition 5.1. *A C^1-curve $\kappa : (0,\infty) \to C$ is called* admissible, *if it satisfies the subsequent conditions:*

(1) *There exists a constant $\delta > 0$ such that $\kappa(t) \in C_\delta$ for all $t > 0$.*
(2) $\lim_{t\to\infty} |\kappa(t)| = \infty$ *and* $\kappa'(t) \in C$ *for all* $t > 0$.



An important class of admissible curves are the rays $\kappa(t) = tx$ with some fixed $x \in C$. In a first step, it is shown that $t \mapsto F(\kappa(t),y)$ is asymptotically constant as $t \to \infty$ for arbitrary admissible curves:

Theorem 5.3. *If $\kappa : (0,\infty) \to C$ is admissible, then for every $y \in C$, the limit*

$$\lim_{t\to\infty} F(\kappa(t),y)$$

exists in $\mathbb{C}^{|G|}$, and is different from 0.

Proof. (Sketch) It is easily calculated that (2.7) translates into the following system of differential equations for $(F_g)_{g\in G}$, where ξ, $y \in \mathbb{R}^N$ are fixed:

$$\partial_\xi F_g(x,y) = \sum_{\alpha\in R_+} k(\alpha)\, \frac{\langle\alpha,\xi\rangle}{\langle\alpha,x\rangle}\, e^{-i\langle\alpha,x\rangle\langle\alpha,gy\rangle} \cdot F_{\sigma_\alpha g}(x,y) \quad (x \in \mathbb{R}^N_{reg}). \quad (5.2)$$

From this, one obtains a differential equation for $x(t) := F(\kappa(t), y)$ of the form $x'(t) = A(t)x(t)$, with a continuous matrix function $A : (0, \infty) \to \mathbb{C}^{|G| \times |G|}$. The proof of Theorem 5.3 is accomplished by verifying that A satisfies the conditions of the following classical theorem on the asymptotic integration of ordinary differential equations (hereby of course, the admissibility conditions on κ come in).

Theorem 5.4. [20], [60]. *Consider the linear differential equation*

$$x'(t) = A(t)x(t), \tag{5.3}$$

where $A : [t_0, \infty) \to \mathbb{C}^{n \times n}$ is a continuous matrix-valued function satisfying the following integrability conditions:

(1) *The matrix-valued improper Riemann integral $\int_{t_0}^{\infty} A(t)dt$ converges.*
(2) *$t \mapsto A(t) \int_t^{\infty} A(s)ds$ belongs to $L^1([t_0, \infty), \mathbb{C}^{n \times n})$.*

Then (5.3) has a fundamental system Φ of solutions which satisfies $\lim_{t \to \infty} \Phi(t) = Id$.

Notice that in the situation of this theorem, for each solution x of $x'(t) = A(t)x(t)$ the limit $\lim_{t \to \infty} x(t)$ exists, and is different from zero, unless $x \equiv 0$.

Proof (of Theorem 5.1). It remains to show that the limit value according to Theorem 5.3 is actually independent of y and κ; the assertion is then easily obtained. The stated independence is accomplished as follows: In a first step, we show that there exists a non-zero vector $v(y) = (v_g(y))_{g \in G} \in \mathbb{C}^{|G|}$ such that for each admissible κ,

$$\lim_{t \to \infty} F(\kappa(t), y) = v(y). \tag{5.4}$$

This can be achieved via interpolation of the admissible curves κ_1, κ_2 by a third admissible curve κ, such that equality of all three limits is being enforced. Next, we focus on admissible rays. Observe that $F_g(tx, y) = F_{g^{-1}}(ty, x)$ for all $g \in G$ and $x, y \in C$. Together with (5.4), this implies that $v_g(y) = v_{g^{-1}}(x)$, and therefore also $v_g(x) = v_{g^{-1}}(x) = v_g(y) =: v_g$. Put $v = (v_g)_{g \in G}$. Then

$$\lim_{t \to \infty} F(\kappa(t), y) = v$$

for every admissible κ and every $y \in C$.

The asymptotic result of Theorem 5.1 also allows to deduce at least a certain amount of information about the structure of the intertwining operator V_k and its representing measures μ_x^k according to formula (4.1). The key for our approach is the following simple observation: according to Corollary 4.1, one may write

$$E_k(x, -i\xi) = \int_{\mathbb{R}^N} e^{-i\langle \xi, y \rangle} d\mu_x^k(y) = \widehat{\mu_x^k}(\xi) \quad (x, \xi \in \mathbb{R}^N),$$

where $\widehat{\mu}$ stands for the (classical) Fourier-Stieltjes transform of $\mu \in M^1(\mathbb{R}^N)$,

$$\widehat{\mu}(\xi) = \int_{\mathbb{R}^N} e^{-i\langle \xi, y\rangle} d\mu(y).$$

Recall that a measure $\mu \in M^1(\mathbb{R}^N)$ is called continuous, if $\mu(\{x\}) = 0$ for all $x \in \mathbb{R}^N$. There is a well-known criterion of Wiener which characterizes Fourier-Stieltjes transforms of continuous measures on locally compact abelian groups, here $(\mathbb{R}^N, +)$; see for instance Lemma 8.3.7 of [22]:

Lemma 5.1. *(Wiener) For $\mu \in M^1(\mathbb{R}^N)$ the following properties are equivalent:*

(1) *μ is continuous.*

(2) $\displaystyle\lim_{n\to\infty} \frac{1}{n^N} \int_{\{\xi\in\mathbb{R}^N:\, |\xi|\leq n\}} |\widehat{\mu}(\xi)|^2 d\xi = 0.$

This yields the following result:

Theorem 5.5. *Let $k \geq 0$. Then apart from the case $k = 0$ (i.e. the classical Fourier case), the measure μ_x^k is continuous for all $x \in \mathbb{R}^N_{reg}$.*

We conclude with two open problems:

(a) In the situation of the last theorem, prove that the measures μ_x^k are even absolutely continuous with respect to Lebesgue measure, provided $\{\alpha \in R : k(\alpha) > 0\}$ spans $\mathbb{R}^N$.

(b) Determine the values of the constants v_g, $g \in G$.

6 Notation

We denote by $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}$ the sets of integer, real and complex numbers respectively. Further, $\mathbb{Z}_+ = \{n \in \mathbb{Z} : n \geq 0\}$. For a locally compact Hausdorff space X, we denote by $C(X), C_b(X), C_c(X)$, $C_0(X)$ the spaces of continuous complex-valued functions on X, those which are bounded, those with compact support, and those which vanish at infinity, respectively. Further, $M_b(X)$, $M_b^+(X)$, $M^1(X)$ are the spaces of regular bounded Borel measures on X, those which are positive, and those which are probability-measures, respectively. Finally, $\mathscr{B}(X)$ stands for the σ-algebra of Borel sets on X.

References

1. Abramowitz, M., Stegun, I.A., *Pocketbook of Mathematical Functions.* Verlag Harri Deutsch, Frankfurt/Main, 1984.
2. Baker, T.H., Forrester, P.J., The Calogero-Sutherland model and generalized classical polynomials. *Comm. Math. Phys.* 188 (1997), 175–216.
3. ———, The Calogero-Sutherland model and polynomials with prescribed symmetry. *Nucl. Phys. B* 492 (1997), 682–716.

4. ______, Non-symmetric Jack polynomials and integral kernels. *Duke Math. J.* 95 (1998), 1–50.
5. Brink, L., Hansson, T.H., Konstein, S., Vasiliev, M.A., The Calogero model - anyonic representation, fermionic extension and supersymmetry. *Nucl. Phys. B* 401 (1993), 591–612.
6. Calogero, F., Solution of the one-dimensional N-body problems with quadratic and/or inversely quadratic pair potentials. *J. Math. Phys.* 12 (1971), 419–436.
7. Chihara, T.S., *An Introduction to Orthogonal Polynomials.* Gordon and Breach, 1978.
8. Davies, E.B.: *Spectral Theory and Differential Operators.* Cambridge University Press, 1995.
9. van Diejen, J.F., Confluent hypergeometric orthogonal polynomials related to the rational quantum Calogero system with harmonic confinement. *Comm. Math. Phys.* 188 (1997), 467–497.
10. van Diejen, J.F., Vinet, L. *Calogero-Sutherland-Moser Models.* CRM Series in Mathematical Physics, Springer-Verlag, 2000.
11. Dunkl, C.F., Reflection groups and orthogonal polynomials on the sphere. *Math. Z.* 197 (1988), 33–60.
12. ______, Differential-difference operators associated to reflection groups. *Trans. Amer. Math. Soc.* 311 (1989), 167–183.
13. ______, Operators commuting with Coxeter group actions on polynomials. In: Stanton, D. (ed.), *Invariant Theory and Tableaux*, Springer, 1990, pp. 107–117.
14. ______, Integral kernels with reflection group invariance. *Canad. J. Math.* 43 (1991), 1213–1227.
15. ______, Hankel transforms associated to finite reflection groups. In: *Proc. of the special session on hypergeometric functions on domains of positivity, Jack polynomials and applications.* Proceedings, Tampa 1991, Contemp. Math. 138 (1992), pp. 123–138.
16. ______, Intertwining operators associated to the group S_3. *Trans. Amer. Math. Soc.* 347 (1995), 3347–3374.
17. ______, Symmetric Functions and B_N-invariant spherical harmonics. Preprint; math.CA/0207122.
18. Dunkl, C.F., de Jeu M.F.E., Opdam, E.M., Singular polynomials for finite reflection groups. *Trans. Amer. Math. Soc.* 346 (1994), 237–256.
19. Dunkl, C.F., Xu, Yuan, *Orthogonal Polynomials of Several Variables;* Cambridge Univ. Press, 2001.
20. Eastham, M.S.P., *The Asymptotic Solution of Linear Differential Systems: Applications of the Levinson Theorem.* Clarendon Press, Oxford 1989.
21. Fell, J.M.G., Doran, R.S., *Representations of $*$-Algebras, Locally Compact Groups , and Banach-$*$-Algebraic Bundles*, Vol. 1. Academic Press, 1988.
22. Graham, C., McGehee, O.C., *Essays in Commutative Harmonic Analysis,* Springer Grundlehren 238, Springer-Verlag, New York 1979.
23. Grove, L.C., Benson, C.T., *Finite Reflection Groups;* Second edition. Springer, 1985.
24. Ha, Z.N.C., Exact dynamical correlation functions of the Calogero-Sutherland model and one dimensional fractional statistics in one dimension: View from an exactly solvable model. *Nucl. Phys. B* 435 (1995), 604–636.
25. Haldane, D., Physics of the ideal fermion gas: Spinons and quantum symmetries of the integrable Haldane-Shastry spin chain. In: A. Okiji, N. Kamakani (eds.),

Correlation effects in low-dimensional electron systems. Springer, 1995, pp. 3–20.
26. Heckman, G.J., A remark on the Dunkl differential-difference operators. In: Barker, W., Sally, P. (eds.) *Harmonic analysis on reductive groups.* Progress in Math. **101**, Birkhäuser, 1991. pp. 181 – 191.
27. ______, Dunkl operators. Séminaire Bourbaki 828, 1996–97; *Astérisque* 245 (1997), 223–246.
28. Helgason, S., *Groups and Geometric Analysis.* American Mathematical Society, 1984.
29. Humphreys, J.E., *Reflection Groups and Coxeter Groups.* Cambridge University Press, 1990.
30. de Jeu, M.F.E., The Dunkl transform. *Invent. Math.* 113 (1993), 147–162.
31. ______, Dunkl operators. Thesis, University of Leiden, 1994.
32. ______, Subspaces with equal closure. Preprint. math.CA/0111015.
33. Kakei, S., Common algebraic structure for the Calogero-Sutherland models. *J. Phys. A* 29 (1996), L619–L624.
34. Kallenberg, O., *Foundations of Modern Probability.* Springer-Verlag, 1997.
35. Knop, F., Sahi, S., A recursion and combinatorial formula for Jack polynomials. *Invent. Math.* 128 (1997), 9–22.
36. Lapointe L., Vinet, L., Exact operator solution of the Calogero-Sutherland model. *Comm. Math. Phys.* 178 (1996), 425–452.
37. Lassalle, M., Polynômes de Laguerre généralisés. *C.R. Acad. Sci. Paris* t. 312 Série I (1991), 725–728
38. ______, Polynômes de Hermite généralisés. *C.R. Acad. Sci. Paris* t. 313 Série I (1991), 579–582.
39. Macdonald, I.G., The Volume of a Compact Lie Group. *Invent. Math.* 56 (1980), 93–95.
40. ______, Some conjectures for root systems. *SIAM J. Math. Anal.* 13 (1982), 988–1007.
41. Moser, J., Three integrable Hamiltonian systems connected with isospectral deformations, *Adv. in Math.* 16 (1975), 197–220.
42. Olshanetsky, M.A., Perelomov, A.M., Completely integrable Hamiltonian systems connected with semisimple Lie algebras. *Invent. Math.* 37 (1976), 93–108.
43. ______, Quantum systems related to root systems, and radial parts of Laplace operators. *Funct. Anal. Appl.* 12 (1978), 121–128.
44. Opdam, E.M., Dunkl operators, Bessel functions and the discriminant of a finite Coxeter group. *Compositio Math.* 85 (1993), 333–373.
45. ______, Harmonic analysis for certain representations of graded Hecke algebras. *Acta Math.* 175 (1995), 75 – 121
46. Pasquier, V.: A lecture on the Calogero-Sutherland models. In: Integrable models and strings (Espoo, 1993), *Lecture Notes in Phys.* 436, Springer, 1994, pp. 36–48.
47. Perelomov, A.M., Algebraical approach to the solution of a one-dimensional model of N interacting particles. *Teor. Mat. Fiz.* 6 (1971), 364–391.
48. Polychronakos, A.P., Exchange operator formalism for integrable systems of particles. *Phys. Rev. Lett.* 69 (1992), 703–705.
49. Rösler, M., Bessel-type signed hypergroups on $\mathbb{R}$. In: Heyer, H., Mukherjea, A. (eds.) Probability measures on groups and related structures XI. Proceedings, Oberwolfach 1994. World Scientific 1995, 292–304.

50. ———, Generalized Hermite polynomials and the heat equation for Dunkl operators. *Comm. Math. Phys.* 192 (1998), 519–542.
51. ———, Positivity of Dunkl's intertwining operator. *Duke Math. J.* 98 (1999), 445–463.
52. ———, Short-time estimates for heat kernels associated with root systems, in *Special Functions,* Conf. Proc. Hong Kong June 1999, eds. C. Dunkl et al., World Scientific, Singapore, 2000, 309–323.
53. ———, One-parameter semigroups related to abstract quantum models of Calogero type. In: *Infinite Dimensional Harmonic Analysis* (Kyoto, Sept. 1999, eds. H. Heyer et al.) Gräbner-Verlag 2000, 290–305.
54. Rösler, M., de Jeu, M., Asymptotic analysis for the Dunkl kernel. math.CA-/0202083; to appear in J. Approx. Theory.
55. Rösler, M., Voit, M., Markov Processes related with Dunkl operators. *Adv. Appl. Math.* 21 (1998), 575–643.
56. Rosenblum, M., Generalized Hermite polynomials and the Bose-like oscillator calculus. In: *Operator Theory: Advances and Applications*, Vol. 73, Basel, Birkhäuser Verlag 1994, 369–396.
57. Sutherland, B., Exact results for a quantum many-body problem in one dimension. *Phys. Rep.* A5 (1972), 1372–1376.
58. Titchmarsh, E.C., *The Theory of Functions.* 2nd ed., Oxford University Press, 1950.
59. Trimèche, K., Paley-Wiener Theorems for the Dunkl transform and Dunkl translation operators. *Integral Transform. Spec. Funct.* 13 (2002), 17–38.
60. A. Wintner, On a theorem of Bôcher in the theory of ordinary linear differential equations, *Amer. J. Math.* 76 (1954), 183–190.
61. Ujino, H., Wadati., M., Rodrigues formula for Hi-Jack symmetric polynomials associated with the quantum Calogero model. *J. Phys. Soc. Japan* 65 (1996), 2423–2439.

Enumeration and Special Functions

Dennis Stanton *

School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455, U.S.A., stanton@math.umn.edu

Summary. These notes for the August 12–16, 2002 Euro Summer School in OPSF at Leuven have three sections with basic introductions to

1. Enumeration and q-series
2. Enumeration and orthogonal polynomials
3. Symmetric functions.

No prior exposure to these areas is assumed. Three excellent textbooks for these three topics are [1], [7], and [17]. Several exercises and **open problems** are given throughout these notes.

1 Enumeration and q-series . . . 138

1.1 q-binomial coefficients . . . 139
1.2 Unimodality . . . 140
1.3 Congruences for the partition function . . . 143
1.4 The Jacobi triple product identity . . . 145
1.5 The Rogers-Ramanujan identities and the involution principle . . . 148
1.6 q-Hermite polynomials and the Rogers-Ramanujan identities . . . 150
1.7 Another q-binomial theorem . . . 153

2 Orthogonal polynomials . . . 155

2.1 General orthogonal polynomials and lattice paths . . . 156
2.2 Hankel determinants . . . 159
2.3 Continued fractions . . . 159

3 Symmetric functions . . . 160

3.1 Combinatorial applications . . . 162
3.2 The Jacobi-Trudi identity . . . 164
3.3 Alternating sign matrices . . . 164
3.4 The Borwein conjecture . . . 165

References . . . 165

* Research partially supported by NSF grant DMS 0203282

E. Koelink and W. Van Assche (Eds.): LNM 1817, pp. 137–166, 2003.

1 Enumeration and q-series

In this section an introduction to q-series and integer partitions is given, with an emphasis on the properties and applications of the q-binomial coefficient.

An *integer partition* is a decreasing sequence $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$ of positive integers. The sum of these integers is the number that λ partitions, $\lambda = 4221$ is a partition of 9. Many generating functions involving integer partitions are q-series, or basic hypergeometric series [14].

As our first example, let a_m be the number of ways to write a non-negative integer m as a sum of distinct integers which are decreasing, namely the number of partitions of m into distinct parts. If $m = 6$, the partitions are

$$6, \quad 51, \quad 42, \quad 321,$$

so $a_6 = 4$. We have

$$\sum_{m=0}^{\infty} a_m q^m = (1+q)(1+q^2)(1+q^3)\cdots = \prod_{j=1}^{\infty}(1+q^j) \tag{1.1}$$

since each part of size j appears exactly once or not at all. A refinement of (1.1) can be made by considering how many parts appear, and writing x^i for a partition with i parts. In our example $a_6(x) = x + 2x^2 + x^3$,

$$\sum_{m=0}^{\infty} a_m(x) q^m = \prod_{j=1}^{\infty}(1+xq^j). \tag{1.2}$$

If we reorganize (1.2) to a power series in x it becomes

$$\sum_{m=0}^{\infty} b_m(q) x^m = \prod_{j=1}^{\infty}(1+xq^j), \tag{1.3}$$

where $b_m(q)$ is the generating function for partitions with exactly m distinct parts. By subtracting one from the mth part, two from the $(m-1)$st part, $\ldots$, up to m from the first part, we obtain a partition where the parts may be equal. In the diagram below (called the *Ferrers diagram* of a partition) we start with $\lambda = 6531$, $m = 4$, and remove 4321.

$$\begin{array}{l} \times\times\times\times\times\times \\ \times\times\times\times\times \\ \times\times\times \\ \times \end{array} \quad \longrightarrow \quad \begin{array}{l} \times\times \\ \times\times \\ \times \\ \end{array}$$

We have removed a triangular array of $\binom{m+1}{2}$ squares; what remains is an arbitrary partition with at most m parts. So

$$b_m(q) = q^{\binom{m+1}{2}} \hat{b}_m(q),$$

where $\hat{b}_m(q)$ is the generating function for all partitions with at most m parts. We may explicitly find $\hat{b}_m(q)$ by considering the conjugate of the Ferrers diagram of a partition with at most m parts: the columns can have any lengths from 1 to m. Letting n_i be the number of times that i occurs and using the geometric series we see that

$$\hat{b}_m(q) = \sum_{n_1,n_2,\cdots,n_m\geq 0} q^{1n_1+2n_2+\cdots+mn_m} = \frac{1}{(1-q)(1-q^2)\cdots(1-q^m)},$$

thus

$$\sum_{m=0}^{\infty} \frac{q^{\binom{m+1}{2}}}{(1-q)\cdots(1-q^m)} x^m = \prod_{j=1}^{\infty}(1+xq^j).$$

1.1 q-binomial coefficients

We may also ask the same question as in the introduction, putting a maximum size N on the partition with distinct parts. The generating function now is

$$\prod_{i=1}^{N}(1+xq^i) = \sum_{m=0}^{\infty} b_m(q,N)x^m, \tag{1.4}$$

where $b_m(q,N)$ is the generating function for partitions with exactly m distinct parts, and largest part at most N. By subtracting one from the mth part, ..., 1 from the largest part as before, we must consider partitions with at most m parts, and largest part at most $N-m$.

Definition 1.1. *The q-binomial coefficient*

$$\begin{bmatrix} N \\ m \end{bmatrix}_q$$

is the generating function for all partitions with at most m parts, and largest part at most $N-m$.

Clearly $\begin{bmatrix} N \\ m \end{bmatrix}_q$ is a polynomial in q of degree $m(N-m)$, whose constant term is 1, corresponding to the empty partition.

We see that (1.4) can be rewritten as (**the q-binomial theorem**)

$$\prod_{i=1}^{N}(1+xq^i) = \sum_{m=0}^{N} \begin{bmatrix} N \\ m \end{bmatrix}_q q^{\binom{m+1}{2}} x^m. \tag{1.5}$$

This is a q-analogue of the binomial theorem, which is the $q \to 1$ limit. In fact there is an explicit "rational formula" for these coefficients.

Proposition 1.1. *The q-binomial coefficient has the explicit formula*

$$\begin{bmatrix} N \\ m \end{bmatrix}_q = \frac{N!_q}{m!_q(N-m)!_q} = \frac{(q;q)_N}{(q;q)_m(q;q)_{N-m}},$$

where

$$n!_q = \prod_{i=1}^{n}(1+q+\cdots+q^{i-1}), \quad (a;q)_n = \prod_{i=0}^{n-1}(1-aq^i).$$

Exercise 1. To verify Proposition 1.1, one can just show that each side satisfies the same recurrence and initial conditions, here the Pascal triangle relation

$$\begin{bmatrix} N \\ m \end{bmatrix}_q = \begin{bmatrix} N-1 \\ m \end{bmatrix}_q + q^{N-m}\begin{bmatrix} N-1 \\ m-1 \end{bmatrix}_q.$$

Can you find another Pascal triangle relation for the q-binomial coefficients?

Exercise 2. Prove the q-binomial theorems in the forms

$$\frac{1}{(1-xq)(1-xq^2)\cdots(1-xq^N)} = \sum_{k=0}^{\infty}\begin{bmatrix} N+k-1 \\ k \end{bmatrix}_q x^k q^k, \tag{1.6}$$

and

$$\frac{(ax;q)_\infty}{(x;q)_\infty} = \sum_{m=0}^{\infty}\frac{(a;q)_m}{(q;q)_m}x^m. \tag{1.7}$$

Exercise 3. Show that the number of vector spaces of dimension m which lie inside a vector space of dimension N over a finite field of order q is

$$\begin{bmatrix} N \\ m \end{bmatrix}_q.$$

Exercise 4. Show that the generating function for all partitions with exactly m parts $(\lambda_1, \lambda_2, \cdots, \lambda_m)$ such that $\lambda_1 - \lambda_{k+1} \le i$ (where $k \le m$) is

$$q^m\frac{(q^{i+1};q)_k}{(q;q)_m}.$$

1.2 Unimodality

The q-binomial coefficient is a polynomial in q whose coefficients are symmetric, and form a *unimodal* sequence. For example,

$$\begin{bmatrix} 7 \\ 3 \end{bmatrix}_q = 1+q+2q^2+3q^3+4q^4+4q^5+5q^6+4q^7+4q^8+3q^9+2q^{10}+q^{11}+q^{12}. \tag{1.8}$$

The symmetry is easy to see, by considering the complement of a Ferrers diagram of a partition inside a rectangle. There is an involved analytic [26] and combinatorial proof [19] of the unimodal property, but it may be shown more easily algebraically. Here is one such proof, which presupposes knowledge of the finite dimensional irreducible representations of the Lie algebra sl_2, the three dimensional algebra of traceless 2×2 matrices, $sl_2 = \text{span}\{e, f, h\}$. (See the contribution of Joris Van der Jeugt in this volume.) We shall use the q-binomial theorem in the proof.

Proof. For each non-negative integer m there is exactly one irreducible representation of sl_2 of dimension $m+1$, denoted $V_m = \{v_{-m}, v_{-m+2}, \cdots, v_m\}$. The basis of V_m may be chosen so that $h(v_i) = iv_i/2$. The *formal character* of V_m, $\text{char}(V_m)$, is the generating function of the dimensions of the h-eigenspaces

$$\begin{aligned}\text{char}(V_m)(q) &= \sum_{\mu} \dim(\mu \text{ eigenspace of } h \text{ in } V_m) q^{\mu}\\ &= q^{-m/2} + q^{-m/2+1} + \cdots + q^{m/2}.\end{aligned}$$

Note that $\text{char}(V_m)(1) = \dim(V_m) = m+1$.

We fix m to be an even positive integer, so that $\text{char}(V_m)$ is a Laurent polynomial in q. We consider the kth exterior power $\wedge^k(V_m)$ of the space V_m, on which sl_2 also acts. This means that $\wedge^k(V_m)$ is the vector space of dimension $\binom{m+1}{k}$ whose basis is

$$v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_k}, \quad \{v_{i_1}, v_{i_2}, \cdots, v_{i_k}\} \subset \{v_{-m}, \cdots, v_m\}.$$

Recall that a Lie algebra acts on a tensor by summing its action on each component. What is the formal character $\text{char}(\wedge^k(V_m))(q)$? If we take all possible k, from 0 to $m+1$, then each basis vector v_i may be chosen once or not at all, just as in our original question about partitions with distinct parts, thus

$$\sum_{k=0}^{m+1} \text{char}(\wedge^k(V_m))(q) t^k = (1+tq^{-m/2})(1+tq^{-m/2+1})\cdots(1+tq^{m/2}).$$

Using the q-binomial theorem we see that

$$\begin{aligned}\text{char}(\wedge^k(V_m))(q) &= q^{-km/2+\binom{k}{2}} \begin{bmatrix} m+1 \\ k \end{bmatrix}_q\\ &= \sum_{i=0}^{k(m+1-k)} c_i q^{-k(m+1-k)/2+i},\end{aligned} \tag{1.9}$$

where

$$\begin{bmatrix} m+1 \\ k \end{bmatrix}_q = \sum_{i=0}^{k(m+1-k)} c_i q^i.$$

Our job is to show that $c_i \leq c_{i+1}$ for $0 \leq i+1 \leq k(m+1-k)/2$. Note that $\mathrm{char}(\wedge^k(V_m))(q)$ is a Laurent polynomial in q since we have taken m to be even.

Now decompose $\wedge^k(V_m)$ into irreducibles,

$$\wedge^k(V_m) = \bigoplus_{s\geq 0} m_s V_s,$$

where m_s is the multiplicity of V_s. We have

$$\mathrm{char}(\wedge^k(V_m))(q) = \sum_{s\geq 0} m_s(q^{-s/2} + q^{-s/2+1} + \cdots + q^{s/2}).$$

Since no half-integer weight occurs in (1.9), we have $c_s = 0$ for s odd. Thus for $i+1 \leq k(m+1-k)/2$,

$$c_i = \sum_{s\geq k(m+1-k)-2i} m_s$$

so that

$$c_{i+1} - c_i = m_{k(m+1-k)-2i-2} \geq 0.$$

As an example of this proof, we have just shown that (1.9) corresponds to the decomposition

$$\wedge^3(V_6) = V_{12} \oplus V_8 \oplus V_6 \oplus V_4 \oplus V_0,$$

$$\begin{aligned} q^{-6}\begin{bmatrix}7\\3\end{bmatrix}_q &= (q^{-6} + q^{-5} + \cdots + q^6) + (q^{-4} + q^{-3} + \cdots + q^4) \\ &\quad + (q^{-3} + q^{-2} + \cdots + q^3) + (q^{-2} + q^{-1} + \cdots + q^2) + 1. \end{aligned}$$

Unimodality is closely related to the following **open problem**, see [23]. Fix an $m \times (N-m)$ rectangle, and consider all partitions whose Ferrers diagrams fit inside this rectangle. Define a partial order $L(m, N-m)$ on these partitions by containment of the respective diagrams. Thus $L(m, N-m)$ has a unique minimal element, the empty partition $\varnothing$, which is covered by the partition 1, which in turn is covered by 11 and 2, until we reach the unique maximum element, the entire rectangle. A *symmetric chain* C in $L(m, N-m)$ is a collection of partitions

$$C = \{\lambda_i, \lambda_{i+1}, \cdots, \lambda_{m(N-m)-i}\}$$

such that λ_j is a partition of j for $i \leq j \leq m(N-m)-i$ and λ_j is covered in $L(m, N-m)$ by λ_{j+1} for $i \leq j < m(N-m)-i-1$. You can imagine C as a saturated chain on the Hasse diagram of $L(m, N-m)$ which is symmetrically located about the middle. A *symmetric chain decomposition* of $L(m, N-m)$ is a collection of disjoint symmetric chains $\{C_s\}$ whose union is

$L(m, N-m)$. Because the q-binomial coefficient is unimodal it is conceivable that $L(m, N-m)$ has such a decomposition. It is unknown if $L(m, N-m)$ has a symmetric chain decomposition for $m \geq 5$, although they have been found for $m \leq 4$.

Exercise 5. Show that if $1 < k < n$, $\gcd(n, k) = 1$, then

$$\frac{1}{[n]_q} \begin{bmatrix} n \\ k \end{bmatrix}_q$$

is a polynomial in q with non-negative coefficients. (**Open problem:** Which partitions does it enumerate?)

1.3 Congruences for the partition function

As another application of the q-binomial theorem we consider $p(n)$, the total number of partitions of n, whose generating function is

$$\sum_{n=0}^{\infty} p(n) q^n = \frac{1}{(q;q)_\infty}.$$

The values of $p(n)$ for $0 \leq n \leq 29$ are tabulated below.

n	$p(n)$	n	$p(n)$
0	1	15	176
1	1	16	231
2	2	17	297
3	3	18	385
4	5	19	490
5	7	20	627
6	11	21	792
7	15	22	1002
8	22	23	1255
9	30	24	1575
10	42	25	1958
11	56	26	2436
12	77	27	3010
13	101	28	3718
14	135	29	4565

Ramanujan proved that

$$p(5n+4) \equiv 0 \mod 5.$$

We now give a proof of this congruence using the q-binomial theorem. The idea is to find a generating function for $p(5n+4)$ with a quadratic form whose symmetry group contains a 5-cycle.

Proof. We start by considering the "sieved" q-binomial theorem: fix non-negative integers M, N, t, r, and s, where $0 \le r < t$, and consider

$$(-xq;q)_{(M+N)t+r} = \prod_{i=1}^{r}(-xq^i;q^t)_{M+N+1} \prod_{i=r+1}^{t}(-xq^i;q^t)_{M+N}. \tag{1.10}$$

Finding the coefficient of x^{Mt+s} in (1.10) we have

$$\begin{bmatrix}(M+N)t+r\\ Mt+s\end{bmatrix}_q q^{\binom{s}{2}} = \sum_{\mathbf{n},\ \mathbf{1}\cdot\mathbf{n}=s} q^{Q(\mathbf{n})} \prod_{i=0}^{r-1}\begin{bmatrix}M+N+1\\ M+n_i\end{bmatrix}_{q^t} \prod_{i=r}^{t-1}\begin{bmatrix}M+N\\ M+n_i\end{bmatrix}_{q^t}, \tag{1.11}$$

where

$$\mathbf{n}=(n_0,\cdots,n_{t-1}),\quad Q(\mathbf{n})=t||\mathbf{n}||^2/2+\mathbf{b}\cdot\mathbf{n}-st/2,\quad \mathbf{b}=(0,1,\cdots,t-1).$$

The $M\to\infty$, $N\to\infty$ limit of (1.11) is

$$\frac{q^{\binom{s}{2}}}{(q;q)_\infty} = \frac{1}{(q^t;q^t)_\infty^t}\sum_{\mathbf{n},\ \mathbf{n}\cdot\mathbf{1}=s} q^{Q(\mathbf{n})}. \tag{1.12}$$

Since

$$Q(\mathbf{n}) = t(n_1^2+\cdots+n_{t-1}^2)+t\sum_{1\le i<j\le t-1} n_in_j + t\binom{s}{2} - ts\sum_{i=1}^{t-1} n_i + \mathbf{b}\cdot\mathbf{n},$$

the r modulo t terms in (1.12) must occur only when $\mathbf{b}\cdot\mathbf{n}\equiv r \mod t$:

$$\sum_{m,m+\binom{s}{2}\equiv r \mod t} p(m)q^{m+\binom{s}{2}} = \frac{1}{(q^t;q^t)_\infty^t}\sum_{\mathbf{n},\mathbf{n}\cdot\mathbf{1}=s,\mathbf{b}\cdot\mathbf{n}\equiv r \mod t} q^{Q(\mathbf{n})}.$$

Choose $t=5$, $s=0$ and $r=4$,

$$\sum_{m=0}^{\infty} p(5m+4)q^{5m+4} = \frac{1}{(q^5;q^5)_\infty^5}\sum_{\mathbf{n},\mathbf{n}\cdot\mathbf{1}=0,\mathbf{b}\cdot\mathbf{n}\equiv 4 \mod 5} q^{Q(\mathbf{n})}.$$

The five $\mathbf{n}$ vectors for $p(4)=5$ are

$$v_0=(1,-1,0,0,0),\quad v_1=(0,1,-1,0,0),\quad v_2=(0,0,1,-1,0),$$
$$v_3=(0,0,0,1,-1),\quad v_4=(1,1,0,-1,-1).$$

If a change of basis is made to $m_0v_0+m_1v_1+m_2v_2+m_3v_3+m_4v_4$, the new quadratic form is

$$\hat{Q}(\mathbf{m}) = 5||\mathbf{m}||^2 - 5(m_0m_1+m_1m_2+m_2m_3+m_3m_4+m_4m_0)-1,$$

with $\mathbf{m}\cdot\mathbf{1}=1$. This clearly has a symmetry group including a 5-cycle, with no fixed points.

This argument also proves $p(7n+5) \equiv 0 \mod 7$ and $p(11n+6) \equiv 0 \mod 11$, see [13]. Dyson proposed a combinatorial proof of Ramanujan's congruence, using the *rank* of a partition,

$$\mathrm{rank}(\lambda) = \text{ largest part of } \lambda - \text{number of parts of } \lambda.$$

It has been proven that the rank modulo 5 splits the partitions of $5n+4$ into 5 equal classes, for example

$$\mathrm{rank}(4) = 3,\ \mathrm{rank}(31) = 1,\ \mathrm{rank}(22) = 0,\ \mathrm{rank}(211) = -1,\ \mathrm{rank}(1111) = -3.$$

It is an **open problem** to find an explicit bijection between these 5 equinumerous rank classes.

Exercise 6. Show that the generating function for all partitions whose rank is $r \geq 0$ is

$$\sum_{m\geq 1} q^{2m+r-1} \begin{bmatrix} 2m+r-2 \\ m-1 \end{bmatrix}_q = \frac{1}{(q;q)_\infty} \sum_{s=1}^{\infty} (-1)^{s-1} q^{s(3s-1)/2+rs}(1-q^s).$$

1.4 The Jacobi triple product identity

One of the most useful results for partitions is the Jacobi triple product identity

$$(-x;q)_\infty(-q/x;q)_\infty(q;q)_\infty = \sum_{m=-\infty}^{\infty} q^{\binom{m}{2}} x^m. \tag{1.13}$$

We sketch three proofs for this result. The first proof again uses the q-binomial theorem,

$$(-x;q)_N(-q/x;q)_N = \sum_{m=-N}^{N} \begin{bmatrix} 2N \\ N+m \end{bmatrix}_q q^{\binom{m}{2}} x^m,$$

and then lets $N \to \infty$ using

$$\lim_{N\to\infty} \begin{bmatrix} 2N \\ N+m \end{bmatrix}_q = \frac{1}{(q;q)_\infty}$$

for any fixed m.

A simple combinatorial proof of (1.13) in the form

$$(-x;q)_\infty(-q/x;q)_\infty = \frac{1}{(q;q)_\infty} \sum_{m=-\infty}^{\infty} q^{\binom{m}{2}} x^m \tag{1.14}$$

is due to Sylvester-Hathaway. The left side of (1.14) is the generating function for pairs of partitions with distinct parts, while the right side is the generating

function for pairs: a triangular partition, and an arbitrary partition. A bijection may be given between these two sets of pairs, by placing a triangle atop a partition, and cutting the resulting diagram it into a pair of partitions with distinct parts. In the example below our triangle 21 has been placed atop 7622, and then cut along the triangle's diagonal to obtain two partitions with distinct parts, 6521 (reading along columns), and 42 (reading along rows).

$$\begin{array}{l} \times \\ \times\ \times \\ \times\ \times\ \times\ \times\ \times\ \times\ \times \\ \times\ \times\ \times\ \times\ \times\ \times \\ \times\ \times \\ \times\ \times \end{array} \longrightarrow \left(\begin{array}{l} \times \\ \times\ \times \\ \times\ \times\ \times \\ \times\ \times\ \times\ \times \\ \times\ \times \\ \times\ \times \end{array} , \begin{array}{l} \times\ \times\ \times\ \times \\ \quad \times\ \times \end{array} \right)$$

Some details need to be checked, but this is the main idea of the proof.

A third proof of the Jacobi triple product identity (1.13) follows by verifying the functional equation $xF(qx) = F(x)$ for

$$F(x) = (-x;q)_\infty(-q/x;q)_\infty = \sum_{n=-\infty}^{\infty} f_n x^n.$$

The functional equation implies that $f_n = q^{n-1} f_{n-1}$, thus $f_n = q^{\binom{n}{2}} f_0$ and we need only find the constant term $f_0 = 1/(q;q)_\infty$ to finish the proof. This may be accomplished combinatorially using the *Frobenius* notation for partitions, or by using the Durfee square in Exercise 7.

Exercise 7. Show that Exercise 2 implies that

$$(-x;q)_\infty = \sum_{k=0}^{\infty} \frac{q^{\binom{k}{2}}}{(q;q)_k}$$

and conclude that

$$f_0 = \sum_{k=0}^{\infty} \frac{q^{k^2}}{(q;q)_k^2}.$$

Finally use the *Durfee square* of a partition—the largest NW justified square in the Ferrers diagram—to conclude that $f_0 = 1/(q;q)_\infty$.

Exercise 8. Use the Jacobi triple product identity (1.13) to find an expansion for $(q;q)_\infty$, and find an involution on partitions with distinct parts which proves this identity. Hopefully you found the Euler Pentagonal Number Theorem.

The *Macdonald identities* [18] generalize the Jacobi triple product identity (1.13) to root systems. The infinite product is now

$$\prod_{\alpha\in\Phi^+} (e^\alpha;q)_\infty(qe^{-\alpha};q)_\infty.$$

In the case of the root system Φ of type A_1, there is one positive root $\Phi^+ = \{\alpha\}$, and if $e^\alpha = -x$, we have $F(x)$. The statement of the theorem is an exact sum expansion for this infinite product. For the root system A_{n-1}, the function to be expanded is

$$F(\mathbf{x}) = F(x_1, \cdots, x_n) = \prod_{1\le i<j\le n} (x_i/x_j;q)_\infty (qx_j/x_i;q)_\infty. \tag{1.15}$$

The third proof of (1.13) also works for the Macdonald identities. First you find a functional equation which reduces the unknown expansion constants to a finite number, then you show that the constant term is sufficient to determine all of these non-zero constants, and finally you evaluate the constant term. For type A_{n-1} the constant term is $1/(q;q)_\infty^{n-1}$. It is an **open problem** to find a combinatorial argument, analogous to either the Frobenius or Durfee square proof for (1.13), which shows that the constant term in $F(\mathbf{x})$ is $1/(q;q)_\infty^{n-1}$.

Here we give the details for the Macdonald identity of type B_2, where the positive roots are $\Phi^+ = \{e_1, e_2, e_2-e_1, e_1+e_2\}$, $e_1 = (1,0)$, $e_2 = (0,1)$. Let

$$F(x,y) = (x, q/x, y, q/y, y/x, qx/y, xy, q/xy; q)_\infty = \sum_{i,j} f_{i,j} x^i y^j.$$

The functional equations are

$$-qx^3 F(qx,y) = F(x,y), \quad -y^3 F(x,qy) = F(x,y),$$

which imply

$$-q^{i-2} f_{i-3,j} = f_{i,j}, \quad -q^{j-3} f_{i,j-3} = f_{i,j},$$

so that a fundamental domain for the constants $f_{i,j}$ is $0 \le i,j \le 2$.

We next use the Weyl group (see Margit Rösler's contribution in this volume), which is generated by reflections in the hyperplanes perpendicular to the roots e_1 and $e_2 - e_1$. These become

$$-xF(1/x,y) = F(x,y), \qquad -f_{1-i,j} = f_{i,j}, \tag{1.16}$$

$$-yF(y,x) = xF(x,y), \qquad -f_{j-1,i} = f_{i-1,j}. \tag{1.17}$$

Equation (1.17) implies that $f_{0,1} = f_{1,2} = f_{2,0} = 0$, along with $f_{0,0} = f_{2,1}$, $f_{1,0} = f_{2,2}$, $f_{1,1} = f_{0,2}$, while (1.16) implies $f_{0,0} = -f_{1,0}$, $f_{0,1} = -f_{1,1}$, $f_{0,2} = -f_{1,0}$. Thus, the only possible non-zero values are taken $f_{0,0} = -f_{1,0} = f_{2,1} = -f_{2,2}$, and we need only evaluate the constant term $f_{0,0}$ in

$$\begin{aligned} &F(x,y) \\ &= f_{0,0} \sum_{i,j} q^{3\binom{j}{2}+3\binom{i}{2}+i} (-1)^{i+j} (x^{3i}y^{3j} - x^{3i+1}y^{3j} + x^{3i+2}y^{3j+1} - x^{3i+2}y^{3j+2}). \end{aligned}$$

Exercise 9. By summing the above identity at $x = 1, \omega, \omega^2$, where $\omega^3 = 1$, show that

$$f_{0,0} = \frac{1}{(q;q)_\infty^2}.$$

The *q-Dyson conjecture* is a finite form of the constant term of the Macdonald identity of type A_{n-1}. Zeilberger and Bressoud [29] proved that the constant term of

$$\prod_{1\le i<j\le n} (x_i/x_j;q)_{a_i}(qx_j/x_i;q)_{a_j}$$

is the q-multinomial coefficient

$$\begin{bmatrix} a_1+\cdots+a_n \\ a_1,\cdots,a_n \end{bmatrix}_q .$$

Their proof uses combinatorial methods similar to the q-series example in Section 1.6. No other proof is known!

1.5 The Rogers-Ramanujan identities and the involution principle

The Rogers-Ramanujan identities are

$$\sum_{n=0}^{\infty} \frac{q^{n^2}}{(q;q)_n} = \frac{1}{(q;q^5)_\infty (q^4;q^5)_\infty}, \tag{1.18}$$

$$\sum_{n=0}^{\infty} \frac{q^{n^2+n}}{(q;q)_n} = \frac{1}{(q^2;q^5)_\infty (q^3;q^5)_\infty}. \tag{1.19}$$

Macmahon found a combinatorial interpretation of these identities.

Proposition 1.2 (Macmahon's Interpretation of RR). *The number of integer partitions of n into parts which differ by at least two is equal to the number of partitions of n into parts congruent to 1 or 4 mod 5. Moreover an analogous result holds if no 1's are allowed for the difference partitions and the mod 5 parts must be 2 or 3.*

If $n=9$, the equinumerous partitions in this statement are

9	9	9	72
81	6111		
72	441	72	333
63	411111	63	3222.
531	111111111		

Schur gave a combinatorial proof of the Rogers-Ramanujan identities in the following form. He considered a generating function for pairs of partitions (λ,μ), where λ has distinct parts and μ has parts which differ by at least two, $(\lambda,\mu)\in \text{Distinct}\times\text{Diff}_2$,

$$(q;q)_\infty \sum_{n=0}^{\infty} \frac{q^{n^2}}{(q;q)_n} = GF(\text{Distinct}\times\text{Diff}_2).$$

The partition λ, which has distinct parts, also has a minus sign attached to each part. Thus the set $\text{Distinct}\times\text{Diff}_2$ may be considered as a "signed" set,

$$\mathrm{sign}((\lambda,\mu)) = (-1)^{\#\text{ parts of }\lambda}.$$

For example,

$$(\varnothing, 94),\ (43, 752) \text{ are positive,} \quad (532, 741),\ (8, 97531) \text{ are negative.}$$

Schur combinatorially defined an involution on the pairs (λ, μ) which changed the number of parts of λ by one, thereby changing its sign, but preserved the number of cells in $\lambda \cup \mu$. His involution cancels all pairs, except for the fixed points, which turn out to be

$$((2p-1, 2p-2, \cdots, p),\ (2p-1, 2p-3, \cdots, 3, 1)), \quad p \geq 0, \tag{1.20}$$
$$((2p, 2p-2, \cdots, p+1),\ (2p-1, 2p-3, \cdots, 3, 1)), \quad p \geq 1. \tag{1.21}$$

Thus Schur proved that

$$(q;q)_\infty \sum_{n=0}^{\infty} \frac{q^{n^2}}{(q;q)_n} = 1 + \sum_{p=1}^{\infty} (-1)^p (q^{p(5p-1)/2} + q^{p(5p+1)/2})$$

and the Jacobi triple product identity (1.13) completes the proof of (1.18).

Schur's clever involution on (λ, μ) was defined in two steps [20].

Step 1 Compare the largest parts λ_1 and μ_1.

1. If $\lambda_1 > \mu_1 + 1$ move λ_1 to μ.
2. If $\mu_1 > \lambda_1$, then move μ_1 to λ.

For example **Step 1** matches $(43, 752) \leftrightarrow (743, 52)$ and $(832, 51) \leftrightarrow (32, 851)$. **Step 1** is not defined if $\lambda_1 = \mu_1$ or if $\lambda_1 = \mu_1 + 1$, and **Step 2** takes care of this possibility.

- Let $lr(\lambda)$ be the length of the leading run of λ. (For example if $\lambda = 87652$, $lr(\lambda) = 4$, from the leading 8765 of λ.)
- Let $sp(\lambda)$ be the smallest part of λ ($sp(87652) = 2$).
- Let $ldr(\mu)$ be the length of the leading double run of μ. (If $\mu = 9752$, then $ldr(\mu) = 3$, from the leading double run 975 of μ.)

Step 2A ($\lambda_1 = \mu_1$)

1. If $sp(\lambda) = \min\{sp(\lambda), lr(\lambda), ldr(\mu)\}$, move the smallest part of λ adjacent to the leading run of λ. (For example $(87652, 8642) \to (9865, 8642)$.)
2. If $ldr(\mu) = \min\{sp(\lambda), lr(\lambda), ldr(\mu)\} < sp(\lambda)$, move the leading double run of μ under the smallest part of λ. (For example $(874, 862) \to (8742, 752)$.)
3. If $lr(\lambda) = \min\{sp(\lambda), lr(\lambda), ldr(\mu)\} < \min\{sp(\lambda), ldr(\mu)\}$, move the leading run of λ to parts $2, 3, \cdots, lr(\lambda) + 1$ of μ, and then move the largest part of μ to a new largest part of λ. (For example, $(984, 9753) \to (9874, 863)$.)

The result of **Step 2A** always gives $\lambda_1 = \mu_1 + 1$.

Step 2B ($\lambda_1 = \mu_1 + 1$)

1. If $lr(\lambda) = \min\{sp(\lambda), lr(\lambda), ldr(\mu)\} < sp(\lambda)$, move the leading run of λ to a new smallest part of λ.

2. If $sp(\lambda) = \min\{sp(\lambda), lr(\lambda), ldr(\mu)\}$, attach the smallest part of λ to the leading double run of μ.
3. If $ldr(\mu) = \min\{sp(\lambda), lr(\lambda), ldr(\mu)\} < \min\{sp(\lambda), lr(\lambda)\}$, move the leading double run of μ to parts $2, 3, \cdots, ldr(\mu) + 1$ of λ, and then move the largest part of λ to a new largest part of μ.

Steps 2A(i) and **2B**(i) are inverses for $i = 1, 2, 3$. It is an exercise to see that **Step 2** is not defined only for the sets (1.20) and (1.21).

While Schur's proof is marvelous, it does not give a direct bijection for Macmahon's interpretation of the Rogers-Ramanujan identities. As of today there is no direct bijection for the Rogers-Ramanujan identities (1.18)–(1.19)! There is an indirect one by Garsia and Milne [12], which uses the *involution principle*.

Here is the general setup for the involution principle. Let $A = A^+ \cup A^-$ be a finite set, consisting of some positive (A^+) and negative (A^-) elements. A *sign-reversing involution* φ is a map $\varphi : A \to A$ such that $\varphi^2 = \varphi$ and if $\varphi(x) \neq x$ then $\mathrm{sign}(x)\mathrm{sign}(\varphi(x)) = -1$. This just means that φ changes sign on its orbits of size 2. Suppose that φ_1 and φ_2 are two sign-reversing involutions on A, with fixed point sets $FP(\varphi_1) \subset A^+$ and $FP(\varphi_2) \subset A^+$ respectively. Note that $|A^+| - |A^-| = |FP(\varphi_1)| = |FP(\varphi_2)|$ so there is a bijection $b : FP(\varphi_1) \to FP(\varphi_2)$. The involution principle guarantees an algorithm to define such a bijection b. If $x \in FP(\varphi_1) \cap FP(\varphi_2)$, then $B(x) = x$, otherwise apply $\varphi_1 \circ \varphi_2$ until you reach $FP(\varphi_2)$. This is guaranteed to occur, since A is finite.

Garsia and Milne showed how to apply this method to Schur's involution [12]. Another involution is necessary to cancel the infinite product $(q; q)_\infty$ when you move it to the other side, thus two involutions are involved.

There are generalizations of the Rogers-Ramanujan identities (1.18)–(1.19) to all moduli involving multisums (see [1]). For example for mod 7 we have

$$\sum_{n_1 \geq n_2 \geq 0} \frac{q^{n_1^2 + n_2^2}}{(q; q)_{n_1 - n_2} (q; q)_{n_2}} = \frac{(q^3, q^4, q^7; q^7)_\infty}{(q; q)_\infty}.$$

1.6 q-Hermite polynomials and the Rogers-Ramanujan identities

There are many proofs of the Rogers-Ramanujan identities [3]. Rogers' original proof used a set of orthogonal polynomials, the continuous q-Hermite polynomials. In this section we give a modern version [24] of Rogers' proof, starting from the well-known integral

$$I(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-xt + t^2/2} e^{-x^2/2}\, dx = e^{t^2}. \tag{1.22}$$

Note that, although (1.22) is easy to prove by completing the square, we are interested in another proof, whose q-analogue will be apparent. The rescaled Hermite polynomials $\hat{H}_n(x) = H_n(x/\sqrt{2})/2^{n/2}$ have the orthogonality relation

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\hat{H}_m(x)\hat{H}_n(x)w(x)\,dx = n!\delta_{mn}, \quad w(x)=e^{-x^2/2}, \tag{1.23}$$

and the generating function

$$G(x,t)=\sum_{n=0}^{\infty}\hat{H}_n(x)\frac{t^n}{n!}=e^{xt-t^2/2}, \tag{1.24}$$

thus (1.22) is the integral of the inverse of the Hermite generating function times the Hermite weight.

We will write down a natural q-analogue of this integral. Evaluating it in two different ways gives the two sides of the Rogers-Ramanujan identities. So we need a q-analogue of the Hermite polynomials, their measure and generating function:

$$H_n(x|q)=\sum_{k=0}^{n}\begin{bmatrix}n\\k\end{bmatrix}_q e^{i(n-2k)\theta}, \quad x=\cos\theta,$$

is a polynomial in $x=\cos\theta$, because it may be rewritten as a sum of Chebyshev polynomials $T_{n-2k}(x)=\cos((n-2k)\theta)$ due to $\begin{bmatrix}n\\k\end{bmatrix}_q=\begin{bmatrix}n\\n-k\end{bmatrix}_q$. These polynomials satisfy the three-term recurrence relation

$$2xH_n(x|q)=H_{n+1}(x|q)-(1-q^n)H_{n-1}(x|q), \tag{1.25}$$

with initial values $H_0(x|q)=1,\ H_1(x|q)=2x$. One can use (1.25) to show that

$$\lim_{q\to 1^-}\frac{H_n(x\sqrt{1-q}/2|q)}{(1-q)^{n/2}}=\hat{H}_n(x).$$

The generating function for $H_n(x|q)$ can easily be found from (1.25):

$$\sum_{n=0}^{\infty}H_n(x|q)\frac{t^n}{(q;q)_n}=\prod_{k=0}^{\infty}(1-2xtq^k+t^2q^{2k})^{-1}=(te^{i\theta},te^{-i\theta};q)_\infty^{-1}.$$

The orthogonality relation for $H_n(x|q)$ is known [14] to be

$$\int_0^{\pi}H_m(\cos\theta|q)H_n(\cos\theta|q)w_q(\cos\theta)\,d\theta=(q;q)_n\delta_{mn},$$

$$w_q(\cos\theta)=\frac{(q;q)_\infty}{2\pi}(e^{2i\theta},e^{-2i\theta};q)_\infty.$$

So the exact analogue of (1.22) for $H_n(x|q)$ is

$$I_q(t)=\frac{(q;q)_\infty}{2\pi}\int_0^{\pi}(te^{i\theta},te^{-i\theta},e^{2i\theta},e^{-2i\theta};q)_\infty\,d\theta.$$

We must evaluate $I_q(t)$ and understand why it is a q-analogue of e^{t^2}. Since we are integrating with respect to the q-Hermite weight, we must find the constant term in the q-Hermite expansion of $(te^{i\theta},te^{-i\theta};q)_\infty$. However

$$(te^{i\theta}, te^{-i\theta}; q)_\infty = \sum_{n=0}^{\infty} H_n(x|q^{-1}) \frac{q^{\binom{n}{2}}(-t)^n}{(q;q)_n},$$

so we just need to find the q-Hermite constant term for $H_n(x|q^{-1})$. Fortunately, Rogers [14] found all of the expansion coefficients

$$H_n(x|q^{-1}) = \sum_{k=0}^{n/2} \frac{q^{k(k-n)}(q;q)_n}{(q;q)_k(q;q)_{n-2k}} H_{n-2k}(x|q),$$

so that

$$\begin{aligned} I_q(t) &= \sum_{n\geq 0,\ \text{even}} \frac{q^{\binom{n}{2}}(-t)^n}{(q;q)_n} \frac{q^{-n^2/4}(q;q)_n}{(q;q)_{n/2}} \\ &= \sum_{m=0}^{\infty} \frac{q^{m^2-m}t^{2m}}{(q;q)_m}. \end{aligned}$$

This is clearly a q-analogue of e^{t^2}, and gives the sum side of (1.18) and (1.19) if $t = \sqrt{q}, q$.

We need an alternative evaluation for $I_q(\sqrt{q})$ and $I_q(q)$ to give the product sides of (1.18) and (1.19). We use exponential orthogonality on $[-\pi, \pi]$ instead of q-Hermite orthogonality. For $t = \sqrt{q}$ the Jacobi triple product identity (1.13) implies

$$(q, \sqrt{q}e^{i\theta}, \sqrt{q}e^{-i\theta}; q)_\infty = \sum_{k=-\infty}^{\infty} (-1)^k q^{k^2/2} e^{-ik\theta},$$

$$(q, e^{2i\theta}, e^{-2i\theta}; q)_\infty = (1-e^{2i\theta}) \sum_{j=-\infty}^{\infty} (-1)^j q^{(j^2+j)/2} e^{2ij\theta},$$

so that

$$\begin{aligned} I_q(\sqrt{q}) &= \frac{1}{2(q;q)_\infty} \sum_{j=-\infty}^{\infty} (q^{2j^2} q^{(j^2+j)/2} - q^{2(j+1)^2} q^{(j^2+j)/2})(-1)^j \\ &= \frac{1}{(q;q)_\infty} \sum_{j=-\infty}^{\infty} q^{2j^2} q^{(j^2+j)/2} (-1)^j = \frac{1}{(q;q^5)_\infty (q^4;q^5)_\infty}. \end{aligned}$$

Note that from the point of view of this proof, the mod 5 condition appears because the q-Hermite weight function involves $e^{2i\theta}$ and $1 + 2^2 = 5$.

Exercise 10. Verify the exponential details to show that $I_q(q)$ does yield (1.19).

1.7 Another q-binomial theorem

We have seen several applications of the q-binomial theorem. Here we give another q-binomial theorem, and shall see it is equivalent to the original q-binomial theorem.

Suppose that x and y are non-commuting variables, but do satisfy $yx = qxy$, where q commutes with x and y. Then any word in x and y can be permuted to x^iy^j at the expense of a power of q, for example, $yxxyyxx = q^8x^4y^3$, because we moved the first y past 4 x's, and the second and third y's past 2 x's.

Theorem 1.1. *If $yx = qxy$ where q commutes with x and y, then*

$$(x+y)^N = \sum_{m=0}^{N} \begin{bmatrix} N \\ m \end{bmatrix}_q x^m y^{N-m}. \tag{1.26}$$

Proof. Each word w of m x's and $N-m$ y's corresponds to a partition which lies inside an $(N-m)\times m$ rectangle. Moreover any such partition corresponds to a unique word w. In our example of $w = yxxyyxx$ the partition is 422. If $w = q^e x^m y^{N-m}$, then e is the sum of the parts of the partition. The generating function for all partitions which lie inside an $(N-m)\times m$ rectangle is the definition of the q-binomial coefficient $\left[\begin{smallmatrix} N \\ N-m \end{smallmatrix}\right]_q = \left[\begin{smallmatrix} N \\ m \end{smallmatrix}\right]_q$.

If we imagine y as a letter larger than x, then (1.26) can be restated as

$$\sum_{\text{words } w \text{ with } m \ x\text{'s},\ N-m \ y\text{'s}} q^{\mathrm{inv}(w)} = \begin{bmatrix} N \\ m \end{bmatrix}_q .$$

where

$$\mathrm{inv}(w) = |\{(i,j) : i<j, \quad w_i > w_j\}|.$$

Exercise 11. Let $W(a_1, a_2, \cdots, a_m) = W(\mathbf{a})$ be the set of all words w of length $a_1 + a_2 + \cdots + a_m$ with exactly a_1 1's, a_2 2's, $\cdots, a_m$ m's. Show that

$$\sum_{w\in W(\mathbf{a})} q^{\mathrm{inv}(w)} = \begin{bmatrix} a_1 + a_2 + \cdots + a_m \\ a_1, a_2, \cdots, a_m \end{bmatrix}_q = \frac{(a_1 + a_2 + \cdots + a_m)!_q}{a_1!_q \, a_2!_q \cdots a_m!_q}.$$

Let's prove the q-analogue of the Pfaff-Saalschütz sum for a 1-balanced ${}_3\varphi_2$ terminating basic hypergeometric sum using Exercise 11 (this proof is due to Zeilberger [27]):

$$\begin{aligned} &\begin{bmatrix} a+b \\ a+k \end{bmatrix}_q \begin{bmatrix} a+c \\ c+k \end{bmatrix}_q \begin{bmatrix} b+c \\ b+k \end{bmatrix}_q \\ &\qquad = \sum_{n=k}^{\min(a,b,c)} q^{n^2-k^2} \frac{[a+b+c-n]!_q}{[a-n]!_q[b-n]!_q[c-n]!_q[n+k]!_q[n-k]!_q}. \end{aligned} \tag{1.27}$$

Proof. Let $S = A_1 \times A_2 \times A_3$ be the set where

1. A_1 consists of all words with $a+k$ 1's, and $b-k$ 2's,
2. A_2 consists of all words with $a-k$ 1's, and $c+k$ 3's,
3. A_3 consists of all words with $b+k$ 2's, and $c-k$ 3's.

We weigh each word by its inversion number, so that the LHS of (1.27) is the generating function of S. We need a bijection from S to words of length $a+b+c-n$ with 5 types of letters, with multiplicities $a-n$, $b-n$, $c-n$, $n+k$ and $n-k$ for some n between k and $min\{a,b,c\}$. We shall see that we may take words in the five symbols

$$\begin{pmatrix}1\\1\\-\end{pmatrix}, \begin{pmatrix}1\\3\\2\end{pmatrix}, \begin{pmatrix}2\\-\\2\end{pmatrix}, \begin{pmatrix}2\\1\\3\end{pmatrix}, \begin{pmatrix}-\\3\\3\end{pmatrix}.$$

The algorithm for the bijection is following. Let $s = (w_1, w_2, w_3) \in S$, and write w_2 under w_1, and w_3 under w_2, all left justified. Scan this "triple word" left to right. If we see

$$\begin{pmatrix}1\\1\\2\end{pmatrix} \text{ or } \begin{pmatrix}1\\1\\3\end{pmatrix}$$

replace it by

$$\begin{pmatrix}1\\1\\-\end{pmatrix},$$

and cross off the 1's in the first 2 rows. Similarly, replace

$$\begin{pmatrix}2\\1\\2\end{pmatrix} \text{ or } \begin{pmatrix}2\\3\\2\end{pmatrix} \text{ by } \begin{pmatrix}2\\-\\2\end{pmatrix},$$

and replace

$$\begin{pmatrix}1\\3\\3\end{pmatrix} \text{ or } \begin{pmatrix}2\\3\\3\end{pmatrix} \text{ by } \begin{pmatrix}-\\3\\3\end{pmatrix}.$$

Keep any $\begin{pmatrix}1\\3\\2\end{pmatrix}$ or $\begin{pmatrix}2\\1\\3\end{pmatrix}$.

For example if $a=3$, $b=4$, $c=3$, $k=1$ and our three words are

$$\begin{aligned} w_1 &= 1\ 2\ 1\ 2\ 2\ 1\ 1 \\ w_2 &= 1\ 3\ 3\ 1\ 3\ 3 \\ w_3 &= 3\ 2\ 2\ 2\ 3\ 2\ 2 \end{aligned}$$

the bijection gives the word in the 5 symbols

$$\begin{pmatrix}1\\1\\-\end{pmatrix}\begin{pmatrix}-\\3\\3\end{pmatrix}\begin{pmatrix}2\\-\\2\end{pmatrix}\begin{pmatrix}1\\3\\2\end{pmatrix}\begin{pmatrix}2\\-\\2\end{pmatrix}\begin{pmatrix}2\\1\\3\end{pmatrix}\begin{pmatrix}1\\3\\2\end{pmatrix}\begin{pmatrix}1\\3\\2\end{pmatrix}.$$

To find out what n is we just solve the equations

$$e_1+e_2=a+k,\ e_1+e_4=a-k,\ e_3+e_4=b-k,\ e_2+e_5=c+k,$$

$$e_2+e_3=b+k,\ e_4+e_5=c-k,$$

to obtain

$$e_1=a-n,\ e_2=n+k,\ e_3=b-n,\ e_4=n-k,\ e_5=c-n,\ \text{for some } n.$$

Our alphabet on the 5 symbols respects inversion, except for $\begin{pmatrix}2\\1\\3\end{pmatrix}\begin{pmatrix}1\\3\\2\end{pmatrix}$, which has 2 inversions instead of 1, thus we must multiply by $q^{e_2e_4}=q^{n^2-k^2}$.

Exercise 12. Show that the *major index*, $\mathrm{maj}(w)$ defined by

$$\mathrm{maj}(w)=\sum_{w_i>w_{i+1}} i$$

also satisfies

$$\sum_{all\ w\ with\ m\ 1's, N-m\ 0's} q^{\mathrm{maj}(w)}=\begin{bmatrix}N\\m\end{bmatrix}_q.$$

Is there a q-multinomial version?

Exercise 13. Give at least two combinatorial proofs (partitions, inv, maj, or finite fields) that

$$\sum_{k=0}^{C}\begin{bmatrix}A\\k\end{bmatrix}_q\begin{bmatrix}B\\C-k\end{bmatrix}_q q^{k(B-C+k)}=\begin{bmatrix}A+B\\C\end{bmatrix}_q.$$

2 Orthogonal polynomials

The classical orthogonal polynomials may be interpreted combinatorially using the exponential formula [9], [10], [11]. One may also give interpretations for general orthogonal polynomials. In this section we consider this general case.

2.1 General orthogonal polynomials and lattice paths

Monic orthogonal polynomials satisfy the three-term recurrence relation

$$p_{n+1}(x) = (x - b_n)p_n(x) - \lambda_n p_{n-1}(x) \tag{2.1}$$

with the initial values $p_{-1}(x) = 0$, $p_0(x) = 1$. Clearly (2.1) implies that $p_n(x)$ is a polynomial in x, with coefficients which are polynomials in the coefficients b_n and λ_n. We give an interpretation for these polynomials using lattice paths in the plane.

Definition 2.1. *A p-lattice path is a lattice path which starts at $(0,0)$, and has steps $NN = (0,2)$, $N = (0,1)$, and $NE = (1,1)$.*

The weight of a p-lattice path, $wt(P)$, is the product of the weights of the individual edges,

$$\begin{aligned} wt((k,j-1) \to (k,j+1)) &= -\lambda_j, \\ wt((k,j) \to (k,j+1) &= -b_j, \\ wt((k,j) \to (k+1,j+1) &= x. \end{aligned}$$

Definition 2.2. *Let* Path_n *be the set of all p-lattice paths which end at $y = n$.*

Proposition 2.1. *The polynomials $p_n(x)$ which satisfy (2.1) with the initial conditions $p_{-1}(x) = 0$ and $p_0(x) = 1$ are given by*

$$p_n(x) = \sum_{P \in \mathrm{Path}_n} wt(P).$$

Even though this proposition is nearly tautological, it is useful. "The" measure $d\mu(x)$ for the polynomials $p_n(x)$ may not be uniquely defined but the moments

$$\mu_n = \int_{-\infty}^{\infty} x^n \, d\mu(x)$$

are, and are also given by a polynomial in b_n and λ_n. (We always assume that $\mu_0 = 1$.) The first few are

$$\begin{aligned} \mu_1 &= b_0 \\ \mu_2 &= b_0^2 + \lambda_1 \\ \mu_3 &= b_0^3 + 2b_0\lambda_1 + b_1\lambda_1 \\ \mu_4 &= b_0^4 + 3b_0^2\lambda_1 + 2b_0b_1\lambda_1 + b_1^2\lambda_1 + \lambda_1^2 + \lambda_1\lambda_2. \end{aligned}$$

Note that (somewhat unexpectedly) all of the coefficients in the moment monomials are non-negative. We shall give a set of lattice paths which are counted by these coefficients.

Definition 2.3. *A Motzkin path is a lattice path which starts at (0,0), lies at or above the x-axis, ends on the x-axis, and has steps* $NE=(1,1)$*,* $E=(1,0)$*, and* $SE=(1,-1)$*.*

Definition 2.4. *Let* Motz_n *be the set of all Motzkin paths from* $(0,0)\to(n,0)$*. The weight of a path is the product of the weights of the individual edges, where*

$$\begin{aligned} wt((k,j)\to(k+1,j+1)) &= 1,\\ wt((k,j)\to(k+1,j)) &= b_j,\\ wt((k,j)\to(k+1,j-1)) &= \lambda_j. \end{aligned}$$

Theorem 2.1 (Viennot [25]). *The* n*th moment* μ_n *is the generating function for Motzkin paths from* $(0,0)\to(n,0)$*,*

$$\mu_n=\sum_{P\in\mathrm{Motz}_n} wt(P).$$

We do not give the proof of Viennot's theorem here. It uses a sign-reversing involution to show that the combinatorial definition of μ_n given above, combined with the combinatorial interpretation of the polynomials $p_n(x)$, does have the appropriate orthogonality relation.

The above proposition and theorem imply that integrals of polynomials (or formal power series) with respect to the the measure $d\mu(x)$ may be evaluated using arguments on lattice paths. To convince you that non-trivial computations can be done with this technique, the Askey-Wilson integral

$$\frac{(q;q)_\infty}{2\pi}\int_0^\pi \frac{(e^{2i\theta},e^{2i\theta};q)_\infty}{(ae^{i\theta},ae^{-i\theta},be^{i\theta},be^{-i\theta},ce^{i\theta},ce^{-i\theta},de^{i\theta},de^{-i\theta};q)_\infty}\,d\theta = \frac{(abcd;q)_\infty}{(ab,ac,ad,bc,bd,cd;q)_\infty}$$

can be evaluated combinatorially [16].

Here are some examples.

Example 2.1. If $b_n=0$ and $\lambda_n=1$, the polynomials $p_n(x)=U_n(x/2)$ (the Chebyshev polynomials of the second kind), and

$$\mu_n=\begin{cases}0, & \text{if } n \text{ is odd},\\ C_N, & \text{if } n=2N \text{ is even}.\end{cases}$$

where $C_n=\frac{1}{n+1}\binom{2n}{n}$ is the nth Catalan number.

Here since $b_n=0$, the Motzkin paths have only NE and SE edges, each with weight one. So μ_n is the total number of such paths, it is well-known [22] that this is a Catalan number.

Example 2.2. If $b_n = 0$ and $\lambda_n = n$, the polynomials $p_n(x) = H_n(x/\sqrt{2})/2^{n/2}$ (the Hermite polynomials), and

$$\mu_n = \begin{cases} 0, & \text{if } n \text{ is odd,} \\ \prod_{i=1}^{N}(2i-1), & \text{if } n = 2N \text{ is even.} \end{cases}$$

Note that μ_n is the number of complete matchings m on n objects, which is the number of ways of matching all n points on the x-axis with $\lfloor n/2 \rfloor$ edges. Here is how to see this. Label the points $1, 2, \cdots, n$. If the kth step of the Motzkin path is NE, then draw an edge emanating from k, to be connected to some point to the right. If the kth step of the Motzkin path is SE, then draw an edge ending at k, to be connected to some earlier point to the left. If the SE edge starts at $y = j$, the weight of the edge is $\lambda_j = j$. This corresponds to the j previous unconnected points, we have j choices for earlier points.

Example 2.3. If $b_n = 2n+\alpha$ and $\lambda_n = n(n+\alpha-1)$, the polynomials $p_n(x) = n!(-1)^n L_n^{\alpha-1}(x)$ (the Laguerre polynomials), and

$$\mu_n = \alpha(\alpha+1)\cdots(\alpha+n-1).$$

If $\alpha = 1$, we have $\mu_n = n!$, the number of permutations of length n. This can be seen by a combinatorial argument not unlike Example 2.2. Here, however, we can do more, by weighing certain choices while building the permutation by α instead of 1, what results are permutation statistics, in this case the number of cycles of a permutation (see [21]).

Example 2.4. If $b_n = n + a$ and $\lambda_n = an$, the polynomials $p_n(x) = C_n^a(x)$ (the Charlier polynomials), and

$$\mu_n = \sum_{k=1}^{n} S(n,k)a^k.$$

This example is similar to Example 2.3. Instead of permutations being in 1-1 correspondence with the weighted Motzkin paths, set partitions are. The Stirling numbers of the second kind $S(n,k)$ is the number of set partitions of $\{1, 2, \cdots, n\}$ into k blocks.

Example 2.5. If $b_n = 0$ and $\lambda_n = [n]_q$, the polynomials
$p_n(x) = H_n(\frac{1}{2}\sqrt{1-q}x|q)/(1-q)^{n/2}$ (the continuous q-Hermite polynomials), and

$$\mu_n = \sum_{\text{complete matchings m on } \{1,2,\cdots,n\}} q^{\mathrm{cross}(m)}.$$

Here $\mathrm{cross}(m)$ is the crossing number of a complete matching m. For example, if $m = (13)(25)(46)$ and we imagine three arcs above the x-axis, these arcs cross twice, $\mathrm{cross}(m) = 2$. In Example 2.2, the j choices now have weights $1,\ q, \cdots, q^{j-1}$. If you connect to the rightmost available previous point, no crossing is produced, as one moves left along the available points, the number of crossings increases by one, thus the choices of $0, 1, \cdots (j-1)$ crossings.

2.2 Hankel determinants

Monic orthogonal polynomials can be defined in terms of the moments μ_n,

$$p_n(x) = \frac{\det \begin{bmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_1 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ 1 & x & \cdots & x^n \end{bmatrix}}{\det \begin{bmatrix} \mu_0 & \mu_1 & \cdots & \mu_{n-1} \\ \mu_1 & \mu_1 & \cdots & \mu_n \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-2} \end{bmatrix}}. \tag{2.2}$$

The coefficient of x^p in the numerator of (2.2) is a minor of the infinite Hankel matrix. Lattice paths give a combinatorial interpretation [25] for any minor.

Theorem 2.2. *The Hankel minor consisting of rows* $\{a_1, a_2, \cdots, a_k\}$ *and columns* $\{b_1, b_2, \cdots, b_k\}$ *(in increasing order) is the generating function for* k*-tuples of non-lattice point intersecting Motzkin paths* $(P_1, \cdots, P_k)$ *such that* $P_i : (-a_i, 0) \to (b_{\sigma(i)}, 0)$ *for some permutation* σ*. The sign of a tuple is given by*

$$\mathrm{sign}((P_1, \cdots, P_k)) = \mathrm{sign}(\sigma).$$

In this theorem if we take rows $(0, 1, 2, \cdots, n-1)$ and columns $(0, 1, 2, \cdots, n-1)$, there is a unique set of paths, and we find the classical fact

$$\lambda_1^{n-1}\lambda_2^{n-2}\cdots\lambda_{n-1}^{1} = \det \begin{bmatrix} \mu_0 & \mu_1 & \cdots & \mu_{n-1} \\ \mu_1 & \mu_1 & \cdots & \mu_n \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-2} \end{bmatrix}.$$

2.3 Continued fractions

The generating function for the moments

$$M(t) = \sum_{n=0}^{\infty} \mu_n t^n$$

is known to have a continued fraction representation

$$M(t) = \cfrac{1}{1 - b_0 t - \cfrac{\lambda_1 t^2}{1 - b_1 t - \cfrac{\lambda_2 t^2}{1 - b_2 t - \cdots}}}. \tag{2.3}$$

Although some care must be taken to consider (2.3) as a complex function of t, as a formal power in t it converges, and we shall see why. This has a very nice combinatorial interpretation which is due to Flajolet [8].

Consider all possible Motzkin paths. Imagine chopping the paths at the line $y=1$. What remains is a sequence of either (a) horizontal edges along $y=0$, or (b) paths starting at $y=1$, ending at $y=1$, along with a precursor NE edge, and a successor SE edge. Since the weight of a horizontal edge along the x-axis is b_0, and the NE-SE pair has weight λ_1, we have

$$M(t)=\sum_{k=0}^{\infty}(b_0t+\lambda_1t^2M^*(t))^k, \tag{2.4}$$

where $M^*(t)$ is the generating function for all Motzkin paths whose weights (b_n and λ_n) have been replaced by (b_{n+1} and λ_{n+1}). Clearly (2.4) is

$$M(t)=\frac{1}{1-b_0t-\lambda_1t^2M^*(t)}. \tag{2.5}$$

Upon iterating (2.5) k times, and then eliminating all paths which go above the line $y=k$, we have the rational function approximation to the continued fraction,

$$M_k(t)=\cfrac{1}{1-b_0t-\cfrac{\lambda_1t^2}{1-b_1t-\cfrac{\lambda_2t^2}{\ddots-\cfrac{\lambda_{k-1}t^2}{1-b_{k-1}t-\lambda_kt^2}}}}. \tag{2.6}$$

Proposition 2.2. *Let $\mu_n(k)$ be the generating function for Motzkin paths of length n which stay at or below $y=k$. Then the generating function*

$$M_k(t)=\sum_{n=0}^{\infty}\mu_n(k)t^n$$

is the kth iterate of the continued fraction (2.3).

It is clear that as formal power series in t,

$$\lim_{k\to\infty}M_k(t)=M(t).$$

Fix any n, then $\mu_n(k)=\mu_n$ for $k\geq n$.

3 Symmetric functions

In this section a quick introduction to symmetric functions is given. The standard reference with a wealth of information is [17].

A polynomial in n variables $x_1, \cdots, x_n$ with complex coefficients which is invariant under all $n!$ permutations may be considered as sum of monomials m_λ, for a partition λ, for example

$$4(x_1^2x_2^1 + x_1^2x_3^1 + x_2^2x_1^1 + x_2^2x_3^1 + x_3^2x_1^1 + x_3^2x_2^1) - 7(x_1^3 + x_2^3 + x_3^3) = 4m_{21} - 7m_3,$$

where

$$\begin{aligned} m_{21}(x_1, x_2, x_3) &= x_1^2x_2^1 + x_1^2x_3^1 + x_2^2x_1^1 + x_2^2x_3^1 + x_3^2x_1^1 + x_3^2x_2^1 \\ m_3(x_1, x_2, x_3) &= x_1^3 + x_2^3 + x_3^3. \end{aligned}$$

The Schur function $s_\lambda(x_1, \cdots, x_n)$ is a symmetric polynomial which may be defined as a quotient of generalized Vandermonde determinants

$$s_\lambda(x_1, \cdots, x_n) = \frac{\det(x_i^{\lambda_j+n-j})_{1\le i,j\le n}}{\det(x_i^{n-j})_{1\le i,j\le n}}. \tag{3.1}$$

Note that the denominator of (3.1) is the Vandermonde determinant

$$\Delta(\mathbf{x}) = \prod_{1\le i<j\le n} (x_i - x_j),$$

and that a quotient of skew-symmetric functions must be symmetric. The Schur functions

$$\{s_\lambda : \lambda \text{ is a partition of } m, \lambda \text{ has at most } n \text{ parts}\}$$

form a basis for the vector space of symmetric polynomials in $x_1, \cdots, x_n$ of degree m.

Special choices for λ give well-known symmetric functions,

$$\begin{aligned} s_{1^k}(\mathbf{x}) &= e_k(\mathbf{x}), \quad \text{the elementary symmetric function of degree } k, \\ s_k(\mathbf{x}) &= h_k(\mathbf{x}), \quad \text{the homogeneous symmetric function of degree } k. \end{aligned}$$

For example

$$\begin{aligned} e_2(x_1, x_2, x_3) &= x_1x_2 + x_1x_3 + x_2x_3, \\ h_2(x_1, x_2, x_3) &= x_1^2 + x_2^2 + x_3^2 + x_1x_2 + +x_1x_3 + x_2x_3 = m_2 + m_{11}. \end{aligned}$$

If we expand the Schur functions in terms of the monomial symmetric functions,

$$s_\lambda(\mathbf{x}) = \sum_\mu K_{\lambda\mu} m_\mu(\mathbf{x}), \tag{3.2}$$

the coefficients $K_{\lambda\mu}$ are called *Kostka* coefficients. These have the algebraic interpretation as the dimension of the weight space μ in an irreducible representation of GL_n that corresponds to λ. Thus they are non-negative integers.

There is also a combinatorial interpretation of $K_{\lambda\mu}$. It is the number of *column strict tableaux* of shape λ and content μ. Take a Ferrers diagram of shape λ, and fill the cells with μ_1 1's, μ_2 2's,..., so that the entries in each row weakly increase and those in each column strictly increase. Here is a column strict tableau of shape $\lambda = 421$ and content $\mu = 2221$.

$$\begin{array}{llll} 1 & 1 & 2 & 4 \\ 2 & 3 & & \\ 3 & & & \end{array}$$

Equation (3.1) is a special case of the Weyl denominator formula for the characters of the general linear group, thus there are analogues of Schur functions and Kostka coefficients for other classical groups.

The characters of the symmetric group S_n may also be found using Schur functions. It is not surprising that such a formula should exist, since the complex irreducible representations of S_n are in 1-1 correspondence with the conjugacy classes of S_n, which are the cycle types, thus partitions of n.

This time we expand the Schur functions in terms of the power sum symmetric functions,

$$p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_m},$$

where

$$p_k = x_1^k + x_2^k + \cdots + x_n^k.$$

For example $p_{21}(x_1, x_2, x_3) = (x_1^2 + x_2^2 + x_3^2)(x_1 + x_2 + x_3)$. We have

$$s_\lambda(\mathbf{x}) = \frac{1}{n!} \sum_\mu c_\mu \chi^\lambda(\mu) p_\mu(\mathbf{x}), \tag{3.3}$$

where c_μ is the the size of the conjugacy class of permutations with cycle type μ,

$$c_\mu = \frac{n!}{\prod_k k^{m_k} m_k!}, \quad \mu = 1^{m_1} 2^{m_2} \cdots,$$

and $\chi^\lambda(\mu)$ is the irreducible character χ^λ evaluated at the conjugacy class μ.

One example of (3.3) is

$$s_{21} = m_{21} + 2m_{111} = \frac{1}{6}(-2p_3 + 2p_{111}),$$

which gives the χ^{21} row of the character table for S_3. Recall that as functions on the symmetric group the characters form an orthonormal basis for the functions constant on conjugacy classes.

3.1 Combinatorial applications

Here I will give two applications of Schur functions to plane partitions (see [17]).

The *hook-content* formula evaluates the principle specialization of a Schur function

$$s_\lambda(1, q, \cdots, q^{n-1}) = q^{n(\lambda)} \prod_{\text{cells } x \in \lambda} \frac{1 - q^{n+c(x)}}{1 - q^{h(x)}}.$$

The *hook* numbers $h(x)$ can be defined by $h(x) = \lambda_i - i + \lambda'_j - j + 1$ if $x = (i, j)$. Pictorially $h(x)$ is the number of cells in the same row or column as x, to the right or below, including x. Here are the hook numbers of each cell for $\lambda = 421$.

$$\begin{matrix} 6 & 4 & 2 & 1 \\ 3 & 1 & & \\ 1 & & & \end{matrix}$$

The *content* numbers $c(x)$ are defined by $c(x) = j - i$ if $x = (i, j)$. Here they are for $\lambda = 421$.

$$\begin{matrix} 0 & 1 & 2 & 3 \\ -1 & 0 & & \\ -2 & & & \end{matrix}$$

The constant $n(\lambda) = \sum_i (i-1)\lambda_i$. The hook-content formula implies

$$h_k(1, q, \cdots, q^{n-1}) = \begin{bmatrix} n+k-1 \\ k \end{bmatrix}_q, \quad e_k(1, q, \cdots, q^{n-1}) = \begin{bmatrix} n \\ k \end{bmatrix}_q q^{\binom{k}{2}},$$

thus the principally specialized Schur function may be considered as a generalization of the q-binomial coefficient. In fact $s_\lambda(1, q, \cdots, q^{n-1})$ is also a symmetric unimodal [17] polynomial in q.

One may also use the hook-content formula to derive the generating function for all *plane partitions*

$$\sum_{n=0}^{\infty} pp(n) q^n = \prod_{k=1}^{\infty} (1 - q^k)^{-k},$$

and the *Macmahon Box Theorem* which gives the generating function for all plane partitions which lie inside an $m \times n \times p$ box

$$\sum_{P \text{ inside } m \times n \times p \text{ box}} q^{\text{size}(P)} = \prod_{1 \le i \le m} \prod_{1 \le j \le n} \prod_{1 \le k \le p} \frac{1 - q^{i+j+k-1}}{1 - q^{i+j+k-2}}.$$

An example of a plane partition which lies inside a $5 \times 3 \times 4$ box is

$$P = \begin{matrix} 4 & 4 & 3 \\ 4 & 2 & \\ 3 & 2 & \\ 1 & & \\ 1 & & \end{matrix}.$$

Here the entries of P lie inside a 5×3 rectangle, and the largest entry is at most 4.

3.2 The Jacobi-Trudi identity

A very useful result, which also has a representation interpretation, is the Jacobi-Trudi identity which gives a Schur function as a single determinant

$$s_\lambda(\mathbf{x}) = \det(h_{\lambda_i - i + j}(\mathbf{x}))_{1 \le i,j \le m},$$

for example

$$s_{421} = \begin{vmatrix} h_4 & h_5 & h_6 \\ h_1 & h_2 & h_3 \\ 0 & h_0 & h_1 \end{vmatrix}.$$

This may also be proven using a "tail-swapping" argument [15] to get non-intersecting lattice paths which correspond to column strict tableaux. This argument is very similar to the involution proofs necessary in Section 2. Moreover realizing plane partitions and column strict tableaux as non-intersecting lattice paths has been a fruitful idea for the study of symmetry classes of plane partitions [6].

3.3 Alternating sign matrices

An $n \times n$ alternating sign matrix is a $0, 1, -1$ matrix whose row and column sums are 1, and whose non-zero entries alternate in sign in every row and column. One example is

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Permutation matrices are examples of alternating sign matrices. Thus alternating sign matrices (ASM) may be considered as generalizations of permutations. A theorem of Zeilberger (see [6] for the whole story) gives the number of $n \times n$ alternating sign matrices as

$$ASM(n) = \prod_{i=0}^{n-1} \frac{(3i+1)!}{(n+i)!}. \tag{3.4}$$

As of today there is no elementary proof of this result! Perhaps even more surprising is that (3.4) also counts the number of totally symmetric self-complementary plane partitions (TSSCPP) which lie inside a $2n \times 2n \times 2n$ box. It is an **open problem** to find a bijection between ASM and TSSCPP, which would allow one to carry over many of the properties of ASM to TSSCPP.

It is remarkable that a generalized determinant, called the λ-determinant [6] may be defined as a sum over alternating sign matrices instead of a sum over permutations. The individual terms are Laurent monomials in the entries of the matrix,

$$\det_\lambda(M) = \sum_{A\in ASM(n)} \lambda^{\mathrm{inv}(A)}(1+\lambda^{-1})^{\#-1's \text{ in } (A)} \prod_{i,j} M_{ij}^{A_{ij}}.$$

Clearly $\det_{-1}(M) = \det(M)$.

3.4 The Borwein conjecture

A final elementary **open problem** is the Borwein conjecture [2, 5]. Let

$$\prod_{i=1}^{n}(1-q^{3i-2})(1-q^{3i-1}) = A_n(q^3) - qB_n(q^3) - q^2C_n(q^3).$$

It is conjectured that the coefficients of the polynomials A_n, B_n, and C_n are non-negative. (For $n=\infty$ it is easy to prove from the Jacobi triple product identity (1.13).) This problem is closely related to hook differences of partitions and statistical physics. Explicit forms for the polynomials may be found (again from the q-binomial theorem), for example

$$A_n(q) = \sum_{k=-n/3}^{n/3} (-1)^k q^{k(9k+1)/2} \begin{bmatrix} 2n \\ n+3k \end{bmatrix}_q,$$

and $A_n(1) = 2\times 3^{n-1}$. Many possible sets of objects can be given which are counted by $A_n(1)$, what is needed is an appropriate statistic. If the quadratic power of q were to change from $k(9k+1)/2$, such a program does work [4].

References

1. G. Andrews, *The Theory of Partitions*, Encyclopedia of Mathematics and Its Applications, Vol. 2 (G.-C. Rota ed.), Addison-Wesley, Reading, Mass., 1976 (reissued by Cambridge Univ. Press, London and New York, 1985).
2. ———, *On a conjecture of Peter Borwein*, J. Symbolic Computation **20** (1995), 487–501.
3. ———, *On the proofs of the Rogers-Ramanujan identities*, in 'q-Series and Partitions', IMA Vol. Math. 18, Springer, New York, 1989, pp. 1–14.
4. G. Andrews, R. Baxter, D. Bressoud, W. Burge, P. Forrester, G. Viennot, *Partitions with prescribed hook differences*, Eur. J. Comb. **8** (1987), 341–350.
5. D. Bressoud, *The Borwein conjecture and partitions with prescribed hook differences*, Elec. J. Comb. **3** (1996), 14 pp.
6. ———, *Proofs and confirmations. The story of the alternating sign matrix conjecture*, Cambridge University Press, Cambridge, 1999.
7. T. S. Chihara, *An Introduction to Orthogonal Polynomials*, Gordon and Breach, New York, 1978.
8. P. Flajolet, *Combinatorial aspects of continued fractions*, Disc. Math. **32** (1980), 125–161.
9. D. Foata, *A combinatorial proof of the Mehler formula*, J. Comb. A **24** (1978), 367–376.

10. D. Foata and P. Leroux, *Polynômes de Jacobi, interpretation combinatoire et fonction génératrice*, Proc. Amer. Math. Soc. **87** (1983), 47–53.
11. D. Foata and V. Strehl, *Combinatorics of Laguerre polynomials*, in 'Enumeration and Design', Academic Press, Toronto, 1984, pp. 123–140.
12. A. Garsia and S. Milne, *Method for constructing bijections for classical partition identities*, Proc. Nat. Acad. Sci. U.S.A. **78** (1981), 2026–2028.
13. F. Garvan, D. Kim, and D. Stanton, *Cranks and t-cores*, Inv. Math. **101** (1990), 1–17.
14. G. Gasper and M. Rahman, *Basic Hypergeometric Series*, Cambridge University Press, Cambridge, 1990.
15. I. Gessel and G. Viennot, *Binomial determinants, paths, and hook length formulae*, Adv. in Math. **58** (1985), 300–321.
16. M. Ismail, D. Stanton, and G. Viennot, *The combinatorics of the q-Hermite polynomials and the Askey-Wilson integral*, Eur. J. Comb. **8** (1987), 379–392.
17. I. G. Macdonald, *Symmetric functions and Hall polynomials*, Oxford University Press, Oxford, 1995.
18. ———, *Affine root systems and Dedekind's η-function*, Inv. Math. **15** (1972), 91–143.
19. K. O'Hara, *Unimodality of Gaussian coefficients: a constructive proof*, J. Comb. A **53** (1990), 29–52.
20. I. Schur, *Ein Beitrag zur additiven Zahlentheorie und zur Theorie der Kettenbrüche*, reprinted in 'I. Schur, Gesammelte Abhandlungen', volume 2, Springer, Berlin, 1973, pp. 117–136.
21. R. Simion and D. Stanton, *Octabasic Laguerre polynomials and permutation statistics*, J. Comput. Appl. Math. **68** (1996), 297–329.
22. R. Stanley, *Enumerative Combinatorics*, Wadsworth, Monterey, 1986.
23. ———, *Weyl groups, the hard Lefschetz theorem, and the Sperner property*, SIAM J. Alg. Disc. Methods **1** (1980), 168–183.
24. D. Stanton, *Gaussian integrals and the Rogers-Ramanujan identities*, in ' Symbolic computation, number theory, special functions, physics, and combinatorics' (F. Garvan and M. Ismail, eds.), Kluwer, Dordrecht, 2001, pp. 255–266.
25. G. Viennot, *Une Théorie Combinatoire des Polynômes Orthogonaux Généraux*, Lecture Notes, University of Quebec at Montreal, 1983.
26. D. Zeilberger, *A one-line high school algebra proof of the unimodality of the Gaussian polynomials $\left[{n \atop k}\right]$ for $k < 20$*, in 'q-Series and Partitions', IMA Vol. Math. 18, Springer, New York, 1989, pp. 67–72.
27. ———, *A q-Foata proof of the q-Saalschütz identity*, Eur. J. Comb. **8** (1987), 461–463.
28. ———, *Proof of the alternating sign matrix conjecture*, Elec. J. Comb. **3** (1996), 1–84.
29. D. Zeilberger and D. Bressoud, *A proof of Andrews' q-Dyson conjecture*, Disc. Math. **54** (1985), 201–224.

Riemann-Hilbert Analysis for Orthogonal Polynomials

Arno B.J. Kuijlaars*

Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200 B, 3001 Leuven, Belgium, arno@wis.kuleuven.ac.be

Summary. This is an introduction to the asymptotic analysis of orthogonal polynomials based on the steepest descent method for Riemann-Hilbert problems of Deift and Zhou. We consider in detail the polynomials that are orthogonal with respect to the modified Jacobi weight $(1-x)^{\alpha}(1+x)^{\beta}h(x)$ on $[-1,1]$ where $\alpha, \beta > -1$ and h is real analytic and positive on $[-1,1]$. These notes are based on joint work with Kenneth McLaughlin, Walter Van Assche and Maarten Vanlessen.

1 **Introduction** ... 168
2 **Boundary values of analytic functions** ... 169
3 **Matrix Riemann-Hilbert problems** ... 172
4 **Riemann–Hilbert problem for orthogonal polynomials on the real line** ... 177
5 **Riemann–Hilbert problem for orthogonal polynomials on $[-1,1]$** ... 179
6 **Basic idea of steepest descent method** ... 181
7 **First transformation $Y \mapsto T$** ... 182
8 **Second transformation $T \mapsto S$** ... 184
9 **Special case $\alpha = \beta = -\frac{1}{2}$** ... 187
10 **Model Riemann Hilbert problem** ... 189
11 **Third transformation $S \mapsto R$** ... 191
12 **Asymptotics for orthogonal polynomials (case $\alpha = \beta = -\frac{1}{2}$)** 192
13 **Case of general α and β** ... 196

* The author's research was supported in part by FWO research project G.0176.02, INTAS project 2000-272, and by the Ministry of Science and Technology (MCYT) of Spain and the European Regional Development Fund (ERDF) through the grant BFM2001-3878-C02-02.

E. Koelink and W. Van Assche (Eds.): LNM 1817, pp. 167–210, 2003.

14 Construction of the local parametrix 198
15 Asymptotics for orthogonal polynomials (general case) 204
References 208

1 Introduction

These lecture notes give an introduction to a recently developed method to obtain asymptotics for orthogonal polynomials. The method is called "steepest descent method for Riemann-Hilbert problems". It is based on a characterization of orthogonal polynomials due to Fokas, Its, and Kitaev [15] in terms of a Riemann-Hilbert problem combined with the steepest descent method introduced by Deift and Zhou in [12] and further developed in [13, 11]. The application to orthogonal polynomials was initiated in the seminal papers of Bleher and Its [4] and Deift, Kriecherbauer, McLaughlin, Venakides and Zhou [8, 9]. These works were motivated by the connection between asymptotics of orthogonal polynomials and universality questions in random matrix theory [7, 27]. An excellent overview can be found in the book of Percy Deift [7], see also [10, 26]. Later developments related to orthogonal polynomials include [2, 3, 5, 18, 20, 21, 22, 23, 34].

In this exposition we will focus on the paper [23] by Kuijlaars, McLaughlin, Van Assche, and Vanlessen. That paper applies the Riemann-Hilbert technique to orthogonal polynomials on the interval $[-1, 1]$ with respect to a modified Jacobi weight $(1-x)^\alpha(1+x)^\beta h(x)$ where h is a non-zero real analytic function on $[-1, 1]$. It should be noted that the earlier works [4, 8, 9] dealt with orthogonal polynomials on the full real line. The fact that one works on a finite interval has some technical advantages and disadvantages. The disadvantage is that one has to pay special attention to the endpoints. The main advantage is that no rescaling is needed on $[-1, 1]$, and that we can work with orthogonal polynomials with respect to a fixed weight function, instead of orthogonal polynomials with respect to varying weights on $\mathbb{R}$. Another advantage is that the analysis simplifies considerably on the interval $[-1, 1]$, if the parameters α and β in the modified Jacobi weight are $\pm\frac{1}{2}$. In that case, there is no need for special endpoint analysis. The case $\alpha = \beta = \frac{1}{2}$ will be worked out in detail in the first part of this paper (up to Section 12).

For general $\alpha, \beta > -1$, the analysis requires the construction of a local parametrix near the endpoints. These local parametrices are built out of modified Bessel functions of order α and β. For orthogonal polynomials on the real line, one typically encounters a parametrix built out of Airy functions [7, 8, 9]. The explicit construction of a local parametrix is a technical (but essential and beautiful) part of the steepest descent method. This is explained in Section 14. The asymptotic behavior of the orthogonal polynomials can then be obtained in any region in the complex plane, including the interval $(-1, 1)$ where the zeros are, and the endpoints ± 1. We will give here the main term in the asymptotic expansion. It is possible to obtain a full asymptotic expansion,

but for this we refer to [23]. We also refer to [23] for the asymptotics of the recurrence coefficients.

We will not discuss here the relation with random matrices. For this we refer to the papers [8] and [24] where the universality for the distribution of eigenvalue spacings was obtained from the Riemann-Hilbert method.

To end this introduction we recall some basic facts from complex analysis that will be used frequently in what follows. First we recall **Cauchy's formula**, which is the basis of all complex analysis. It says that

$$f(z) = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(s)}{s-z} ds$$

whenever f is analytic in a domain Ω, γ is a simple closed, positively oriented curve in Ω, encircling a domain Ω_0 which also belongs to Ω, and $z \in \Omega_0$.

A second basic fact is **Liouville's theorem**, which says that a bounded entire function is constant. An extension of Liouville's theorem is the following. If f is entire and $f(z) = \mathcal{O}(z^n)$ as $z \to \infty$, then f is a polynomial of degree at most n.

Exercise 1. If you have not seen this extension of Liouville's theorem before, (or if you forgot about it) try to prove it.

We also recall **Morera's theorem**, which says that if f is continuous in a domain Ω and satisfies $\oint_\gamma f(z)dz = 0$ for all closed contours γ in Ω, then f is analytic in Ω.

We will also use some basic facts about isolated singularities of analytic functions. In a basic course in complex analysis you learn that an isolated singularity is either removable, a pole, or an essential singularity. **Riemann's theorem on removable singularities** says that if an analytic function is bounded near an isolated singularity, then the singularity is removable. The following is an extension of this result.

Exercise 2. Let $a \in \Omega$. If f is analytic in $\Omega \setminus \{a\}$, and $\lim_{z\to a}(z-a)f(z) = 0$, then a is a removable singularity of f.

2 Boundary values of analytic functions

We will deal with boundary values of analytic functions on curves. Suppose γ is a curve, which could be an arc, or a closed contour, or a system of arcs and contours. We will always consider **oriented curves**. The orientation induces a $+$side and a $-$side on γ. By definition, the $+$side is on the left, while traversing γ according to its orientation, and the $-$side is on the right.

All curves we consider are smooth (C^1 or even analytic), but the curves may have points of self-intersection or endpoints. At such points the $+$ and $-$sides are not defined. We use γ^o to denote γ without points of self-intersection and endpoints.

Let f be an analytic function on $\mathbb{C} \setminus \gamma$. The boundary values of f in $s \in \gamma^o$ are defined by

$$f_+(s) = \lim_{\substack{z \to s \\ z \text{ on } +\text{side}}} f(z), \qquad f_-(s) = \lim_{\substack{z \to s \\ z \text{ on } -\text{side}}} f(z),$$

provided these limits exist. If these limits exist for every $s \in \gamma^o$, and f_+ and f_- are continuous functions on γ^o, then we say that f has **continuous boundary values** on γ^o. It is possible to study boundary values in other senses, like L^p-sense, see [7, 14], but here we will always consider boundary values in the sense of continuous boundary values.

If the boundary values f_+ and f_- of f exists, and if we put $v(s) = f_+(s) - f_-(s)$, we see that f satisfies the following **Riemann-Hilbert problem** (boundary value problem for analytic functions)

(RH1) f is analytic in $\mathbb{C} \setminus \gamma$.
(RH2) $f_+(s) = f_-(s) + v(s)$ for $s \in \gamma^o$.

We say that v is the jump for f over γ^o.

Suppose now that conversely, we are given $v(s)$ for $s \in \gamma^o$. Then we may ask ourselves whether the above Riemann-Hilbert problem has a solution f, and whether the solution is unique. It is easy to see that the solution cannot be unique, since we can add an entire function to f and obtain another solution. So we need to impose an extra condition to guarantee uniqueness. This is typically an asymptotic condition, such as

(RH3) $f(z) \to 0$ as $z \to \infty$.

In this way we have normalized the Riemann-Hilbert problem at infinity. It is also possible to normalize at other points, but we will only meet problems where the normalization is at infinity.

It turns out that there is a unique solution if v is Hölder continuous on γ and if γ is a simple closed contour, or a finite disjoint union of simple closed contours, see [17, 28]. Then there are no points of self-intersection or endpoints so that $\gamma = \gamma^o$. In the case of points of self-intersection or endpoints, we need extra conditions at those points.

If γ is a simple closed contour, oriented positively, and if v is Hölder continuous on γ, then it can be shown that

$$f(z) = \frac{1}{2\pi i} \oint_\gamma \frac{v(s)}{s - z} ds \tag{2.1}$$

is the unique solution of the Riemann-Hilbert problem (RH1), (RH2), (RH3). It is called the **Cauchy transform** of v and we denote it also by $C(v)$. We will not go into the general existence theory here (for general theory, see e.g. [17, 28]), but leave it as an (easy) exercise for the case of analytic v.

Exercise 3. Assume that v is analytic in a domain Ω and that γ is a simple closed contour in Ω. Prove that the Cauchy transform of v satisfies the Riemann-Hilbert problem (RH1), (RH2), (RH3).
[Hint: Use a deformation of γ and Cauchy's formula.]

Exercise 4. Give an explicit solution of the Riemann-Hilbert problem (RH1), (RH2), (RH3), for the case where the jump v is a rational function with no poles on the simple closed contour γ.
[Hint: Use partial fraction decomposition of v.]

To establish uniqueness, one assumes as usual that there are two solutions f_1 and f_2. Then the difference $g = f_1 - f_2$ will solve a homogeneous Riemann-Hilbert problem with trivial jump $g_+ = g_-$ on γ. Then it follows from Morera's theorem that g is analytic on γ. Hence g is an entire function. From the asymptotic condition (RH3) it follows that $g(z) \to 0$ as $z \to \infty$, and therefore, by Liouville's theorem, g is identically zero.

We will use the above argument, based on Morera's theorem, also in other situations. We leave it as an exercise.

Exercise 5. Suppose that γ is a curve, or a system of curves, and that f is analytic on $\mathbb{C} \setminus \gamma$. Let γ_0 be an open subarc of γ^o so that f has continuous boundary values f_+ and f_- on γ_0 that satisfy $f_+ = f_-$ on γ_0. Show that f is analytic across γ_0.

In the case that γ has points of self-intersection or endpoints, extra conditions are necessary at the points of $\gamma \setminus \gamma^o$. We consider this for the case of the interval $[-1, 1]$ in the following exercise.

Exercise 6.

(a) Suppose that γ is the interval $[-1, 1]$ and v is a continuous function on $(-1, 1)$. Also assume that the Riemann-Hilbert problem (RH1), (RH2), (RH3) has a solution. Show that the solution is not unique.
(b) Show that there is a unique solution if we impose, in addition to (RH1), (RH2), (RH3), the conditions that

$$\lim_{z \to 1} (z-1) f(z) = 0, \qquad \lim_{z \to -1} (z+1) f(z) = 0.$$

The next step is to go from an additive Riemann-Hilbert problem to a multiplicative one. This means that instead of (RH2), we have a jump condition

(RH4) $f_+(s) = f_-(s) v(s)$ for $s \in \gamma^o$.

In this case, the asymptotic condition is typically

(RH5) $f(z) \to 1$ as $z \to \infty$.

If γ is a simple closed contour, and if v is continuous and non-zero on γ, then we define the index (or winding number) of v by

$$\operatorname{ind} v = \frac{1}{2\pi} \Delta_\gamma v(s).$$

This is $\frac{1}{2\pi}$ times the change in the argument of $v(s)$ as we go along γ once in the positive direction. The index is an integer. If the index is zero, then we

can take a continuous branch of the logarithm of v on γ, and obtain from (RH4) the additive jump condition

$$(\log f)_+(s) = (\log f)_-(s) + \log v(s), \qquad s \in \gamma.$$

This has a solution as a Cauchy transform

$$\log f(z) = \frac{1}{2\pi i} \oint_\gamma \frac{\log v(s)}{s-z} ds,$$

provided $\log v$ is Hölder continuous on γ. Then

$$f(z) = \exp\left(\frac{1}{2\pi i} \oint_\gamma \frac{\log v(s)}{s-z} ds \right) \tag{2.2}$$

solves the additive Riemann-Hilbert problem (RH1), (RH4), (RH5).

Exercise 7. How would you solve a Riemann-Hilbert problem with jump condition

$$f_+(s) f_-(s) = v(s)$$

for s on a simple closed contour γ?

3 Matrix Riemann-Hilbert problems

The Riemann-Hilbert problems that are associated with orthogonal polynomials are stated in a matrix form for 2×2 matrix valued analytic functions.

A matrix valued function $R : \mathbb{C} \setminus \gamma \to \mathbb{C}^{2\times 2}$ is analytic if all four entries of R are analytic functions on $\mathbb{C} \setminus \gamma$. Then a typical Riemann-Hilbert problem for 2×2 matrices is the following

(mRH1) $R : \mathbb{C} \setminus \gamma \to \mathbb{C}^{2\times 2}$ is analytic.
(mRH2) $R_+(s) = R_-(s) V(s)$ for $s \in \gamma^o$, where $V : \gamma^o \to \mathbb{C}^{2\times 2}$ is a given matrix valued function on γ^o.
(mRH3) $R(z) \to I$ as $z \to \infty$, where I denotes the 2×2 identity matrix.

Because of (mRH3) the problem is normalized at infinity. The matrix valued function V in (mRH2) is called the **jump matrix**. If $\gamma^o \neq \gamma$, then additional conditions have to be imposed at the points of self-intersection and the endpoints.

The existence theory of the **matrix Riemann-Hilbert problem** given by (mRH1), (mRH2), (mRH3) is quite complicated, and we will not deal with it. In the problems that we will meet, we know from the beginning that there is a solution. We know the solution because it is built out of orthogonal polynomials. See [6] for a systematic treatment of the general theory of matrix Riemann-Hilbert problems.

Also, we will only meet Riemann-Hilbert problems where the jump matrix V satisfies

$$\det V(s) = 1, \qquad s \in \gamma^o. \tag{3.1}$$

Then we can establish uniqueness of the Riemann-Hilbert problem (mRH1), (mRH2), (mRH3) on a simple closed contour. The argument is as follows.

First we consider the scalar function $\det R : \mathbb{C} \setminus \gamma \to \mathbb{C}$. It is analytic on $\mathbb{C} \setminus \gamma$, and in view of (mRH2) and (3.1) we have $(\det R)_+ = (\det R)_-$ on γ. Thus $\det R$ is an entire function. From (mRH3) it follows that $\det R(z) \to 1$ as $z \to \infty$, so that by Liouville's theorem $\det R(z) = 1$ for every $z \in \mathbb{C} \setminus \gamma$. Now suppose that $\tilde{R}$ is another solution of (mRH1), (mRH2), (mRH3). Since $R(z)$ has determinant 1, we can take the inverse, and we consider $X(z) = \tilde{R}(z)[R(z)]^{-1}$. Then X is clearly analytic on $\mathbb{C} \setminus \gamma$, and for $s \in \gamma$,

$$\begin{aligned} X_+(s) = \tilde{R}_+(s)[R_+(s)]^{-1} &= \tilde{R}_-(s)V(s)[R_-(s)V(s)]^{-1} \\ &= \tilde{R}_-(s)R_-(s) = X_-(s). \end{aligned}$$

Thus X is entire. Finally, we have $X(z) \to I$ as $z \to \infty$, so that by Liouville's theorem, we get $X(z) = I$ for every $z \in \mathbb{C} \setminus \gamma$, which shows that $\tilde{R} = R$. Thus R is the unique solution of (mRH1), (mRH2), (mRH3).

While existence of the solution is not an issue for us, we do need a result on the behavior of the solution R. We need to know that when V is close to the identity matrix on γ, then R is close to the identity matrix in the complex plane. We will need this result only for simple closed contours γ and it will be enough for us to deal with jump matrices that are analytic in a neighborhood of γ.

In order to specify the notion of closeness to the identity matrix, we need a norm on matrices. We can take any matrix norm, but for definiteness we will take the matrix infinity norm (maximum row sum) defined for 2×2 matrices R by

$$\|R\| = \max\left(|R_{11}| + |R_{12}|, |R_{21}| + |R_{22}|\right).$$

If $R(z)$ is a matrix-valued function defined on a set Ω, we define

$$\|R\|_\Omega = \sup_{z \in \Omega} \|R(z)\|,$$

where for $\|R(z)\|$ we use the infinity norm. If $R(z)$ is analytic on a domain Ω, then one may show that $\|R(z)\|$ is subharmonic as a function of z. If $R(z)$ is also continuous on $\bar{\Omega}$, then by the the maximum principle for subharmonic functions, it assumes its maximum value on the boundary of Ω.

With these preliminaries we can establish the following result. The following elementary complex analysis proof is due to A.I. Aptekarev [2].

Theorem 3.1. *Suppose γ is a positively oriented simple closed contour and Ω is an open neighborhood of γ. Then there exist constants C and $\delta > 0$ such that a solution R of the matrix Riemann-Hilbert problem (mRH1), (mRH2), (mRH3) with a jump matrix V that is analytic on Ω with*

$$\|V - I\|_\Omega < \delta,$$

satisfies

$$\|R(z) - I\| < C\|V - I\|_\Omega \tag{3.2}$$

for every $z \in \mathbb{C} \setminus \gamma$.

Proof. In the proof we use $\mathrm{ext}(\gamma)$ and $\mathrm{int}(\gamma)$ to denote the exterior and interior of γ, respectively. So, $\mathrm{ext}(\gamma)$ is the unbounded component of $\mathbb{C} \setminus \gamma$, and $\mathrm{int}(\gamma)$ is the bounded component. Together with γ, we also consider two simple closed curves γ_e and γ_i, both homotopic to γ in Ω, so that $\gamma_e \subset \Omega \cap \mathrm{ext}(\gamma)$ and $\gamma_i \subset \Omega \cap \mathrm{int}(\gamma)$, see Figure 1.

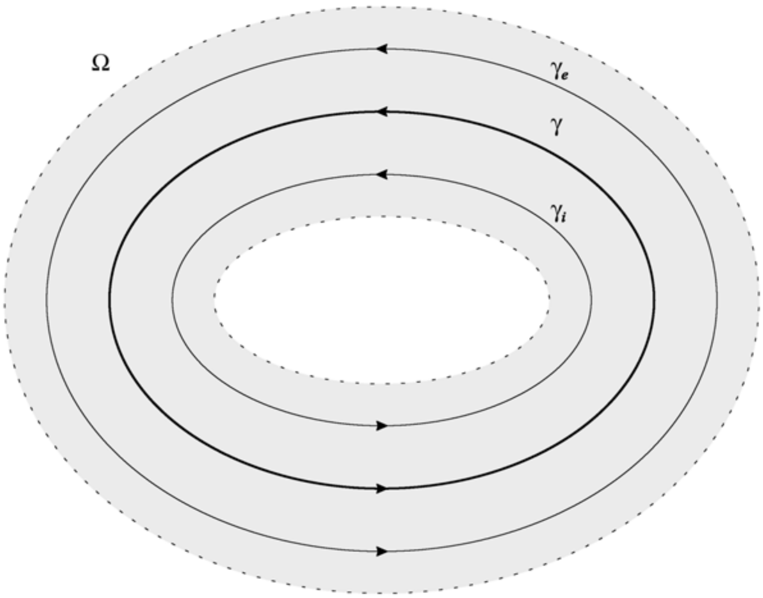


Fig. 1. Illustration for the proof of Theorem 3.1. The shaded region is the domain Ω, which contains the simple closed curves γ, γ_e, and γ_i.

We choose $r > 0$ so that

$$\min\left(\mathrm{dist}(z, \gamma_e), \mathrm{dist}(z, \gamma_i)\right) > r \qquad \text{for every } z \in \gamma, \tag{3.3}$$

where $\mathrm{dist}(z, \gamma_e)$ and $\mathrm{dist}(z, \gamma_i)$ denote the distances from z to the respective curves.

We write $\Delta = V - I$. Since $R_+ = R_- + R_-\Delta$, we may view $R_-\Delta$ as an additive jump for R on γ. By (2.1) and the asymptotic condition (mRH3) we thus have

$$R(z) = I + \frac{1}{2\pi i} \oint_\gamma \frac{R_-(s)\Delta(s)}{s - z} ds \tag{3.4}$$

for $z \in \mathbb{C} \setminus \gamma$. The integral in (3.4) is taken entrywise.

The idea of the proof is to show that for $\|\Delta\|_\Omega$ small enough, we have $\|R_-(s)\| \leq 4$ for every $s \in \gamma$. (Any other positive number than 4 would also do.) If we can prove this, then it follows by straightforward estimation on (3.4) that

$$\begin{aligned} \|R(z) - I\| &= \left\| \frac{1}{2\pi i} \oint_\gamma \frac{R_-(s)\Delta(s)}{s-z} ds \right\| \\ &\leq \frac{4l(\gamma)}{2\pi \mathrm{dist}(z,\gamma)} \|\Delta\|_\gamma \\ &\leq \frac{4l(\gamma)}{2\pi \mathrm{dist}(z,\gamma)} \|V - I\|_\Omega \end{aligned} \tag{3.5}$$

where $l(\gamma)$ is the length of γ. This then proves (3.2) for $\mathrm{dist}(z,\gamma) > r$ with constant

$$C = \frac{4l(\gamma)}{2\pi r}.$$

To handle the case when z is close to γ, we apply the same arguments to the curves γ_e and γ_i. Suppose for example that $z \in \mathrm{ext}(\gamma)$. Then we define

$$\tilde{R} = \begin{cases} R & \text{in } \mathrm{ext}(\gamma) \cup \mathrm{int}(\gamma_i) \\ RV^{-1} & \text{in } \mathrm{int}(\gamma) \cap \mathrm{ext}(\gamma_i) \end{cases}$$

Then $\tilde{R}_+ = \tilde{R}_-$ on γ so that $\tilde{R}$ is analytic across γ. On γ_i we have the jump $\tilde{R}_+ = \tilde{R}_- V$. The same arguments we will give below that lead to $\|R_-(s)\| \leq 4$ for $s \in \gamma$ will also show that $\|\tilde{R}_-(s)\| \leq 4$ for $s \in \gamma_i$ (provided $\|\Delta\|_\Omega$ is sufficiently small). Then an estimate similar to (3.5) shows that for every z,

$$\|\tilde{R}(z) - I\| \leq \frac{4l(\gamma_i)}{2\pi \mathrm{dist}(z,\gamma_i)} \|V - I\|_\Omega.$$

For $z \in \mathrm{ext}(\gamma)$, we have $\tilde{R}(z) = R(z)$ and $\mathrm{dist}(z,\gamma_i) > r$ by (3.3), so that we get (3.2) with a maybe different constant C. The same arguments apply for $z \in \mathrm{int}(\gamma)$. In that case we define $\tilde{R}$ so that it has a jump on γ_e.

So it remains to prove that $\|R_-(z)\| \leq 4$ for every $z \in \gamma$. In order to do this we put

$$M = \max_{z \in \gamma} \|R_-(z)\|.$$

Since $R_-(z)$ are the continuous boundary values for R taken from $\mathrm{ext}(\gamma)$, and R is analytic in $\mathrm{ext}(\gamma)$, including the point at infinity, we have by the maximum principle for subharmonic functions, that

$$\|R(z)\| \leq M, \qquad z \in \mathrm{ext}(\gamma).$$

We deform γ to γ_e lying in $\Omega \cap \mathrm{ext}(\gamma)$. Then $\mathrm{dist}(z,\gamma_e) > r$ for every $z \in \gamma$ by (3.3). For $z \in \mathrm{int}(\gamma)$, we then have

$$R(z) = I + \frac{1}{2\pi i} \oint_{\gamma_e} \frac{R(s)\Delta(s)}{s-z} ds.$$

Letting z go to γ from within $\mathrm{int}(\gamma)$, we then find

$$R_+(z) = I + \frac{1}{2\pi i} \oint_{\gamma_e} \frac{R(s)\Delta(s)}{s-z} ds, \qquad z \in \gamma,$$

and so, since $R_+ = R_-(I + \Delta)$,

$$R_-(z) = \left(I + \frac{1}{2\pi i} \oint_{\gamma_e} \frac{R(s)\Delta(s)}{s-z} ds\right) (I + \Delta(z))^{-1}, \qquad z \in \gamma.$$

We take norms, and estimate, where we use that $\|R(s)\| \leq M$ for $s \in \gamma_e$,

$$\|R_-(z)\| \leq \left(1 + \frac{l(\gamma_e)}{2\pi r} M \|\Delta\|_\Omega\right) \left\|(I + \Delta(z))^{-1}\right\|, \qquad z \in \gamma,$$

with $l(\gamma_e)$ the length of γ_e. If $\|\Delta(z)\| \leq \frac{1}{2}$ then $\left\|(I + \Delta(z))^{-1}\right\| \leq 1 + 2\|\Delta(z)\|$, which follows easily from estimating the Neumann series

$$(I + \Delta(z))^{-1} = \sum_{k=0}^{\infty} (-\Delta(z))^k.$$

So we assume $\delta > 0$ is small enough so that

$$\delta < \frac{1}{2} \qquad \text{and} \qquad \frac{l(\gamma_e)}{2\pi r} \delta(1 + 2\delta) < \frac{1}{2}.$$

Then, if $\|\Delta\|_\Omega < \delta$, we find for $z \in \gamma$,

$$\begin{aligned} \|R_-(z)\| &\leq \left(1 + \frac{l(\gamma_e)}{2\pi r} M\delta\right)(1 + 2\delta) \\ &= (1 + 2\delta) + \frac{l(\gamma_e)}{2\pi r} M\delta(1 + 2\delta) \\ &\leq 2 + \frac{1}{2}M. \end{aligned}$$

Taking the supremum for $z \in \gamma$, we get $M \leq 2 + \frac{1}{2}M$, which means that $M \leq 4$. So we have proved our claim that $\|R_-(z)\| \leq 4$ for every $z \in \gamma$, which completes the proof of the theorem.

Exercise 8. Analyze the proof of Theorem 3.1 and show that we can strengthen (3.2) to

$$\|R(z) - I\| \leq \frac{C}{1 + |z|} \|V - I\|_\Omega$$

for every $z \in \mathbb{C} \setminus \gamma$.

4 Riemann–Hilbert problem for orthogonal polynomials on the real line

Fokas, Its, and Kitaev [15] found a characterization of orthogonal polynomials in terms of a matrix Riemann-Hilbert problem.

We consider a weight function w on $\mathbb{R}$, which is smooth and has sufficient decay at $\pm\infty$, so that all moments $\int x^k w(x)\,dx$ exist. The weight induces a scalar product $\int f(x)g(x)w(x)dx$, and the Gram-Schmidt orthogonalization process applied to the sequence of monomials $1, x, x^2, \ldots,$ yields a sequence of **orthogonal polynomials** $\pi_0, \pi_1, \pi_2, \ldots$, that satisfy

$$\int \pi_n(x)\pi_m(x)w(x)\,dx = h_n\delta_{n,m}, \qquad h_n > 0.$$

We will choose the polynomials to be monic $\pi_n(x) = x^n + \cdots$. If we put

$$\gamma_n = h_n^{-1/2}, \qquad p_n(x) = \gamma_n\pi_n(x)$$

then the polynomials p_n are the orthonormal polynomials, i.e.,

$$\int p_n(x)p_m(x)w(x)dx = \delta_{n,m}.$$

The orthonormal polynomials satisfy a **three-term recurrence**

$$xp_n(x) = a_{n+1}p_{n+1}(x) + b_np_n(x) + a_np_{n-1}(x)$$

with certain recurrence coefficients a_n and b_n. The monic form of the recurrence is

$$x\pi_n(x) = \pi_{n+1}(x) + b_n\pi_n(x) + a_n^2\pi_{n-1}(x).$$

Fokas, Its, Kitaev [15] formulated the following Riemann-Hilbert problem for a 2×2 matrix valued function $Y : \mathbb{C}\setminus\mathbb{R} \to \mathbb{C}^{2\times 2}$.

(Y-RH1) Y is analytic in $\mathbb{C}\setminus\mathbb{R}$.

(Y-RH2) $Y_+(x) = Y_-(x)\begin{pmatrix} 1 & w(x) \\ 0 & 1 \end{pmatrix}$ for $x \in \mathbb{R}$.

(Y-RH3) $Y(z) = \left(I + \mathcal{O}\left(\frac{1}{z}\right)\right)\begin{pmatrix} z^n & 0 \\ 0 & z^{-n} \end{pmatrix}$ as $z \to \infty$.

The asymptotic condition (Y-RH3) does not say that $Y(z)$ tends to the identity matrix as z tends to infinity (unless $n = 0$), so the problem is not normalized at infinity in the sense of (mRH3).

Theorem 4.1 (Fokas, Its, Kitaev). *The Riemann-Hilbert problem (Y-RH1)–(Y-RH3) for Y has a unique solution given by*

$$Y(z) = \begin{pmatrix} \pi_n(z) & C(\pi_n w)(z) \\ c_n\pi_{n-1}(z) & c_nC(\pi_{n-1}w)(z) \end{pmatrix} \tag{4.1}$$

*where π_n and π_{n-1} are the **monic** orthogonal polynomials of degrees n and $n-1$, respectively, $C(\pi_j w)$ is the Cauchy transform of $\pi_j w$,*

$$C(\pi_j w)(z) = \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{\pi_j(x)w(x)}{x-z} dx,$$

and c_n is the constant

$$c_n = -2\pi i \gamma_{n-1}^2.$$

Proof. Consider the first row of Y. The condition (Y-RH2) gives for the first entry Y_{11}

$$(Y_{11})_+(x) = (Y_{11})_-(x), \qquad x \in \mathbb{R}.$$

Thus Y_{11} is an entire function. The asymptotic condition (Y-RH3) gives

$$Y_{11}(z) = z^n + \mathcal{O}(z^{n-1}) \qquad \text{as } z \to \infty.$$

By the extension of Liouville's theorem, this implies that Y_{11} is a monic polynomial of degree n. We call it P_n.

Now we look at Y_{12}. The jump condition (Y-RH2) gives

$$(Y_{12})_+(x) = (Y_{12})_-(x) + (Y_{11})_-(x)w(x).$$

We know already that $Y_{11} = P_n$, so that

$$(Y_{12})_+(x) = (Y_{12})_-(x) + P_n(x)w(x). \tag{4.2}$$

The asymptotic condition (Y-RH3) implies

$$Y_{12}(z) = \mathcal{O}\left(z^{-n-1}\right) \qquad \text{as } z \to \infty. \tag{4.3}$$

The conditions (4.2)–(4.3) constitute an additive scalar Riemann-Hilbert problem for Y_{12}. Its solution is given by the Cauchy transform

$$Y_{12}(z) = C(P_n w)(z) = \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{P_n(x)w(x)}{x-z} dx.$$

Now in general the Cauchy transform tends to zero like z^{-1} as $z \to \infty$, and not like z^{-n-1} as required in (4.3). We need extra conditions on the polynomial P_n to ensure that (4.3) is satisfied. We write

$$\frac{1}{x-z} = -\sum_{k=0}^{n-1} \frac{x^k}{z^{k+1}} + \frac{x^n}{z^n(x-z)}.$$

Then

$$\begin{aligned} Y_{12}(z) &= \frac{1}{2\pi i} \int P_n(x)w(x) \left[-\sum_{k=0}^{n-1} \frac{x^k}{z^{k+1}} + \frac{x^n}{z^n(x-z)} \right] dx \\ &= -\sum_{k=0}^{n-1} \frac{1}{2\pi i} \left[\int P_n(x)x^k w(x)\, dx \right] \frac{1}{z^{k+1}} + \mathcal{O}\left(z^{-n-1}\right). \end{aligned}$$

In order to have (4.3) we need that the coefficient of z^{-k-1} vanishes for $k = 0, \dots, n-1$. Thus

$$\int P_n(x) x^k w(x)\, dx = 0, \qquad k = 0, \dots, n-1.$$

This means that P_n is the orthogonal polynomial, and since P_n is monic, we have $P_n = \pi_n$. Thus we have shown that the first row of Y is equal to the expressions given in the equality (4.1). The equality for the second row is shown in a similar way. The details are left as an exercise.

Exercise 9. Show that the second row of Y is equal to the expressions given in (4.1).

Remark concerning the proof of Theorem 4.1 The above proof of Theorem 4.1 is not fully rigorous in two respects. First, we did not check that the jump condition (Y-RH2) is valid in the sense of continuous boundary values, and second, we did not check that the asymptotic condition (Y-RH3) holds uniformly as $z \to \infty$ in $\mathbb{C} \setminus \mathbb{R}$. This is not immediate since $\mathbb{R}$ is an unbounded contour.

Both of these questions are technical issues whose treatment falls outside the scope of this introduction. Suitable smoothness and decay properties have to be imposed on w. The reader is referred to [9, Appendix A] for a discussion of these matters. There it is shown that it is enough that $x^n w(x)$ belongs to the Sobolev space H^1 for every n.

5 Riemann–Hilbert problem for orthogonal polynomials on $[-1, 1]$

We will study polynomials that are orthogonal with respect to weights on the finite interval $[-1, 1]$. In particular we will consider modified Jacobi weight

$$w(x) = (1-x)^\alpha (1+x)^\beta h(x), \qquad x \in (-1, 1) \tag{5.1}$$

where $\alpha, \beta > -1$ and h is positive on $[-1, 1]$ and analytic in a neighborhood of $[-1, 1]$. The weights (5.1) are a generalization of the Jacobi weights which have $h(x) \equiv 1$. In analogy with the case of the whole real line, the Riemann-Hilbert problem that characterizes the orthogonal polynomials has the following ingredients.

We look for a matrix valued function $Y : \mathbb{C} \setminus [-1, 1] \to \mathbb{C}^{2\times 2}$ that satisfies

(Y-RH1) Y is analytic in $\mathbb{C} \setminus [-1, 1]$.

(Y-RH2) $Y_+(x) = Y_-(x) \begin{pmatrix} 1 & w(x) \\ 0 & 1 \end{pmatrix}$ for $x \in (-1, 1)$

(Y-RH3) $Y(z) = \left(I + \mathcal{O}\left(\frac{1}{z}\right) \right) \begin{pmatrix} z^n & 0 \\ 0 & z^{-n} \end{pmatrix}$ as $z \to \infty$

Note that we restrict ourselves in the jump condition (Y-RH2) to the open interval $(-1,1)$. The jump is not defined at the endpoints ± 1, since the boundary values $Y_\pm$ are not defined there. If α or β (or both) is negative, there is also a problem with the definition of w at the endpoints.

We can show, as for the case of orthogonal polynomials on the real line, that

$$Y(z) = \begin{pmatrix} \pi_n(z) & C(\pi_n w)(z) \\ c_n \pi_{n-1}(z) & c_n C(\pi_{n-1} w)(z) \end{pmatrix} \tag{5.2}$$

is a solution of the Riemann-Hilbert problem, where now C denotes the Cauchy transform on $[-1,1]$, that is,

$$C(\pi_n w)(z) = \frac{1}{2\pi i} \int_{-1}^{1} \frac{\pi_n(x) w(x)}{x - z}\, dx.$$

However, this will not be the only solution. In order to ensure uniqueness we need extra conditions at the endpoints ± 1. The endpoint conditions are

(Y-RH4) As $z \to 1$, we have

$$Y(z) = \begin{cases} \mathcal{O}\begin{pmatrix} 1 & |z-1|^\alpha \\ 1 & |z-1|^\alpha \end{pmatrix} & \text{if } \alpha < 0, \\[2ex] \mathcal{O}\begin{pmatrix} 1 & \log|z-1| \\ 1 & \log|z-1| \end{pmatrix} & \text{if } \alpha = 0, \\[2ex] \mathcal{O}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} & \text{if } \alpha > 0. \end{cases}$$

(Y-RH5) As $z \to -1$, we have

$$Y(z) = \begin{cases} \mathcal{O}\begin{pmatrix} 1 & |z+1|^\beta \\ 1 & |z+1|^\beta \end{pmatrix} & \text{if } \beta < 0, \\[2ex] \mathcal{O}\begin{pmatrix} 1 & \log|z+1| \\ 1 & \log|z+1| \end{pmatrix} & \text{if } \beta = 0, \\[2ex] \mathcal{O}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} & \text{if } \beta > 0. \end{cases}$$

In (Y-RH4)–(Y-RH5) the $\mathcal{O}$ conditions are to be taken entrywise, so the condition (Y-RH4) in the case $\alpha < 0$ means that

$$\begin{array}{ll} Y_{11}(z) = \mathcal{O}(1) & \quad Y_{12}(z) = \mathcal{O}(|z-1|^\alpha) \\ Y_{21}(z) = \mathcal{O}(1) & \quad Y_{22}(z) = \mathcal{O}(|z-1|^\alpha) \end{array}$$

as $z \to 1$. So Y_{11} and Y_{21} should remain bounded at $z = 1$, while Y_{12} and Y_{22} are allowed to grow as $z \to 1$, but not faster than $\mathcal{O}(|z-1|^\alpha)$.

Now we can prove that Y given by (5.2) satisfies the boundary conditions (Y-RH4)–(Y-RH5), and that it is in fact the only solution to the Riemann-Hilbert problem (Y-RH1)–(Y-RH5). This is left to the reader as an exercise (see also [23]).

Exercise 10. Show that (5.2) satisfies the conditions (Y-RH4)–(Y-RH5).

Exercise 11. Show that the Riemann-Hilbert problem (Y-RH1)-(Y-RH5) for Y has (5.2) as its unique solution.

6 Basic idea of steepest descent method

The steepest descent method for Riemann-Hilbert problems consists of a sequence of explicit transformations, which in our case have the form

$$Y \mapsto T \mapsto S \mapsto R.$$

The ultimate goal is to arrive at a Riemann-Hilbert problem for R on a system of contours γ,

(R-RH1) R is analytic on $\mathbb{C} \setminus \gamma$,
(R-RH2) $R_+(s) = R_-(s)V(s)$ for $s \in \gamma$,
(R-RH3) $R(z) \to I$ as $z \to \infty$,

in which the jump matrix V is close to the identity.

Note that Y depends on n through the asymptotic condition

$$Y(z) = \left(I + \mathcal{O}\left(\frac{1}{z}\right)\right) \begin{pmatrix} z^n & 0 \\ 0 & z^{-n} \end{pmatrix},$$

and so, to indicate the n-dependence, we may write $Y = Y^{(n)}$. Also the transformed functions T, S, and R depend on n, say $T = T^{(n)}$, $S = S^{(n)}$, and $R = R^{(n)}$. The jump matrix $V = V^{(n)}$ in (R-RH2) also depends on n. The contour γ, however, does not depend on n. The jump matrices that we will find have analytic continuations to a neighborhood of γ, which is also independent of n, and we will have

$$V^{(n)}(s) = I + \mathcal{O}\left(\frac{1}{n}\right) \qquad \text{as } n \to \infty,$$

uniformly for s in a neighborhood of γ. Then, from Theorem 3.1, we can conclude that

$$R^{(n)}(z) = I + \mathcal{O}\left(\frac{1}{n}\right) \qquad \text{as } n \to \infty,$$

uniformly for $z \in \mathbb{C} \setminus \gamma$. Tracing back the steps $Y^{(n)} \mapsto T^{(n)} \mapsto S^{(n)} \mapsto R^{(n)}$, we find asymptotics for $Y^{(n)}$, valid uniformly in the complex plane. So, in particular, since π_n is the $(1,1)$ entry of Y_{11}, we find asymptotic formulas for the orthogonal polynomials that are uniformly valid in every region of the complex plane.

The steepest descent method for Riemann-Hilbert methods is an alternative to more classical asymptotic methods that have been developed for differential equations or integral representations. The Jacobi polynomials $P_n^{(\alpha,\beta)}$

that are orthogonal with respect to $(1-x)^\alpha(1+x)^\beta$ have an integral representation and they satisfy a second order differential equation. As a result their asymptotic behavior as $n \to \infty$ is very well-known, see [31]. The orthogonal polynomials associated with the weights (5.1) do not have an integral representation or a differential equation, and so asymptotic methods that are based on these cannot be applied. The steepest descent method for Riemann-Hilbert problems is the first method that is able to give full asymptotic expansions for orthogonal polynomials in a number of cases where integral representations and differential equations are not available.

It must be noted that other methods, based on potential theory and approximation theory, have also been used for asymptotics of orthogonal polynomials [25, 30, 32]. These methods apply to weights with less smoothness, but the results are not as strong as the ones we will present here.

7 First transformation $Y \mapsto T$

The first transformation uses the mapping

$$\varphi(z) = z + (z^2-1)^{1/2}, \qquad z \in \mathbb{C} \setminus [-1,1].$$

That branch of the square root is chosen which is analytic in $\mathbb{C} \setminus [-1,1]$ and which is positive for $z > 1$. Thus $(z^2-1)^{1/2}$ is negative for real $z < -1$.

Exercise 12. Show the following

(a) φ is a one-to-one map from $\mathbb{C}\setminus[-1,1]$ onto the exterior of the unit disk.
(b) $\varphi(z) = 2z + \mathcal{O}\left(\frac{1}{z}\right)$ as $z \to \infty$.
(c) $\varphi_+(x)\varphi_-(x) = 1$ for $x \in (-1,1)$.

The first transformation is

$$T(z) = \begin{pmatrix} 2^n & 0 \\ 0 & 2^{-n} \end{pmatrix} Y(z) \begin{pmatrix} \varphi(z)^{-n} & 0 \\ 0 & \varphi(z)^n \end{pmatrix}. \tag{7.1}$$

Then straightforward calculations show that T satisfies the Riemann-Hilbert problem

(T-RH1) T is analytic in $\mathbb{C} \setminus [-1,1]$.
(T-RH2) $T_+(x) = T_-(x) \begin{pmatrix} \varphi_+(x)^{-2n} & w(x) \\ 0 & \varphi_-(x)^{-2n} \end{pmatrix}$ for $x \in (-1,1)$.
(T-RH3) $T(z) = I + \mathcal{O}\left(\frac{1}{z}\right)$ as $z \to \infty$.
(T-RH4)–(T-RH5) T has the same behavior as Y near ± 1.

Exercise 13. Verify that the jump condition (T-RH2) and the asymptotic condition (T-RH3) hold.

The effect of the transformation $Y \mapsto T$ is that the problem is normalized at infinity, since $T(z) \to I$ as $z \to \infty$. This is good. What is not so good, is

that the jump matrix for T is more complicated. The entries on the diagonal have absolute value one, and so for large n, they are rapidly oscillating as x varies over the interval $(-1,1)$. The effect of the next transformation will be to transform these oscillations into exponentially small terms.

Why did we choose to perform the transformation (7.1)? An easy answer would be: because we will see later that it works. A second answer could be based on a list of desirable properties that the function φ should have. The honest answer is that already a lot is known about orthogonal polynomials and their asymptotics, see, e.g., [16, 25, 29, 31, 33]. For example it is known that

$$\lim_{n\to\infty} (\pi_n(z))^{1/n} = \frac{\varphi(z)}{2}, \qquad z \in \mathbb{C} \setminus [-1,1] \tag{7.2}$$

where that branch of the nth root is chosen which behaves like z at infinity. This is the **nth root asymptotics** of the polynomials π_n. It is intimately connected with the weak convergence of zeros. The nth root asymptotics (7.2) holds for a very large class of weights w on $(-1,1)$. It is for example enough that $w > 0$ almost everywhere on $(-1,1)$.

A stronger kind of asymptotics is

$$\lim_{n\to\infty} \frac{2^n \pi_n(z)}{\varphi(z)^n} = \frac{\tilde{D}(\infty)}{\tilde{D}(z)}, \tag{7.3}$$

which is valid uniformly for $z \in \mathbb{C} \setminus [-1,1]$. The **strong asymptotics** (7.3) is valid for weights w satisfying the **Szegő condition**, that is,

$$\int_{-1}^{1} \frac{\log w(t)}{\sqrt{1-t^2}} dt > -\infty.$$

The function $\tilde{D}$ appearing in the right-hand side of (7.3) is known as the **Szegő function**. It is analytic and non-zero on $\mathbb{C} \setminus [-1,1]$, and there it is a finite limit

$$\lim_{z\to\infty} \tilde{D}(z) = \tilde{D}(\infty) \in (0,\infty).$$

Note that in what follows, we will use a different definition for the Szegő function, and we will call it D, instead of $\tilde{D}$.

Since we want to recover the asymptotics (7.2)–(7.3) (and more) for the modified Jacobi weights, we cannot avoid using the functions that appear there. This explains why we perform the transformation (7.1). The $(1,1)$ entry of T is

$$T_{11}(z) = \frac{2^n \pi_n(z)}{\varphi(z)^n}$$

and this is a quantity which we like. We have peeled off the main part of the asymptotics of π_n. By (7.3) we know that the limit of $T_{11}(z)$ exists as $n \to \infty$, and the limit is expressed in terms of the Szegő function associated with w. This indicates that the transformation $Y \mapsto T$ makes sense. It also indicates that we will have to use the Szegő function in one of our future transformations.

Exercise 14. Another idea would be to define

$$\tilde{T}(z) = Y(z) \begin{pmatrix} \varphi(z)^{-n} & 0 \\ 0 & \varphi(z)^n \end{pmatrix} \begin{pmatrix} 2^n & 0 \\ 0 & 2^{-n} \end{pmatrix}.$$

This would also lead to the $(1,1)$ entry being $\frac{2^n \pi_n(z)}{\varphi(z)}$. Work out the Riemann-Hilbert problem for $\tilde{T}$. What is the advantage of T over $\tilde{T}$?

Exercise 15. The transformation

$$\widehat{T}(z) = \begin{pmatrix} \varphi(z)^{-n} & 0 \\ 0 & \varphi(z)^n \end{pmatrix} Y(z) \begin{pmatrix} 2^n & 0 \\ 0 & 2^{-n} \end{pmatrix}$$

would lead to the same $(1,1)$ entry, but this transformation is a very bad idea. Why?

8 Second transformation $T \mapsto S$

The second transformation $T \mapsto S$ is based on a factorization of the jump matrix in (T-RH2)

$$\begin{pmatrix} \varphi_+^{-2n} & w \\ 0 & \varphi_-^{-2n} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{1}{w}\varphi_-^{-2n} & 1 \end{pmatrix} \begin{pmatrix} 0 & w \\ -\frac{1}{w} & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{1}{w}\varphi_+^{-2n} & 1 \end{pmatrix},$$

which can be verified by direct calculation.

Instead of making one jump across the interval $(-1,1)$, we can now think that we are making three jumps according to the above factorization. That is, if we cross the interval $(-1,1)$ from the upper half-plane into the lower half-plane, we will first make the jump $\begin{pmatrix} 1 & 0 \\ \frac{1}{w}\varphi_+^{-2n} & 1 \end{pmatrix}$, then the jump $\begin{pmatrix} 0 & w \\ -\frac{1}{w} & 0 \end{pmatrix}$, and finally the jump $\begin{pmatrix} 1 & 0 \\ \frac{1}{w}\varphi_-^{-2n} & 1 \end{pmatrix}$.

Now recall that $w(x) = (1-x)^\alpha(1+x)^\beta h(x)$ is the modified Jacobi weight. The extra factor h is positive on $[-1,1]$ and analytic in a neighborhood of $[-1,1]$. Then there is a neighborhood U of $[-1,1]$ so that h is analytic on U with positive real part, see Figure 2. All our future deformations will be contained in U.

We will consider $(1-z)^\alpha$ as an analytic function on $\mathbb{C} \setminus [1,\infty)$ where we take the branch which is positive for real $z < 1$. Similarly, we will view $(1+z)^\beta$ as an analytic function on $\mathbb{C} \setminus (-\infty,-1]$. Then

$$w(z) = (1-z)^\alpha(1+z)^\beta h(z)$$

is non-zero and analytic on $U \setminus ((-\infty,-1] \cup [1,\infty))$, and it is an analytic continuation of our weight $w(x)$.

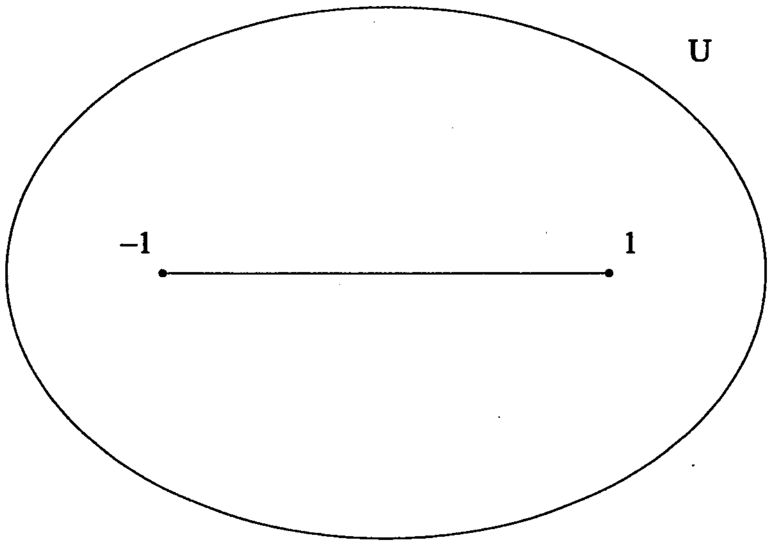


Fig. 2. Neighborhood U of $[-1,1]$ so that h is analytic with positive real part in U.

The two jump matrices $\begin{pmatrix} 1 & 0 \\ \frac{1}{w}\varphi_+^{-2n} & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ \frac{1}{w}\varphi_-^{-2n} & 1 \end{pmatrix}$ then have natural extensions into the upper and lower half-planes, respectively, both given by $\begin{pmatrix} 1 & 0 \\ \frac{1}{w}\varphi^{-2n} & 1 \end{pmatrix}$. Note that for z away from the interval $[-1,1]$, we have $|\varphi(z)| > 1$, so that $\begin{pmatrix} 1 & 0 \\ \frac{1}{w}\varphi(z)^{-2n} & 1 \end{pmatrix}$ is close to the identity matrix if n is large.

We open a lens-shaped region around $(-1,1)$ as shown in Figure 3 of the paper. The lens is assumed to be contained in the domain U. The upper and lower lips of the lens are denoted by Σ_1 and Σ_3 respectively. The interval $[-1,1]$ is denoted here by Σ_2.

Then we define the second transformation $T \mapsto S$ by

$$S = \begin{cases} T & \text{outside the lens} \\ T\begin{pmatrix} 1 & 0 \\ -\frac{1}{w}\varphi^{-2n} & 1 \end{pmatrix} & \text{in the upper part of the lens} \\ T\begin{pmatrix} 1 & 0 \\ \frac{1}{w}\varphi^{-2n} & 1 \end{pmatrix} & \text{in the lower part of the lens.} \end{cases} \tag{8.1}$$

The transformation results in jumps for S on the interior of the three curves Σ_1, $\Sigma_2 = [-1,1]$ and Σ_3. It follows that S satisfies the following Riemann-Hilbert problem

(S-RH1) S is analytic in $\mathbb{C} \setminus (\Sigma_1 \cup [-1,1] \cup \Sigma_3)$.

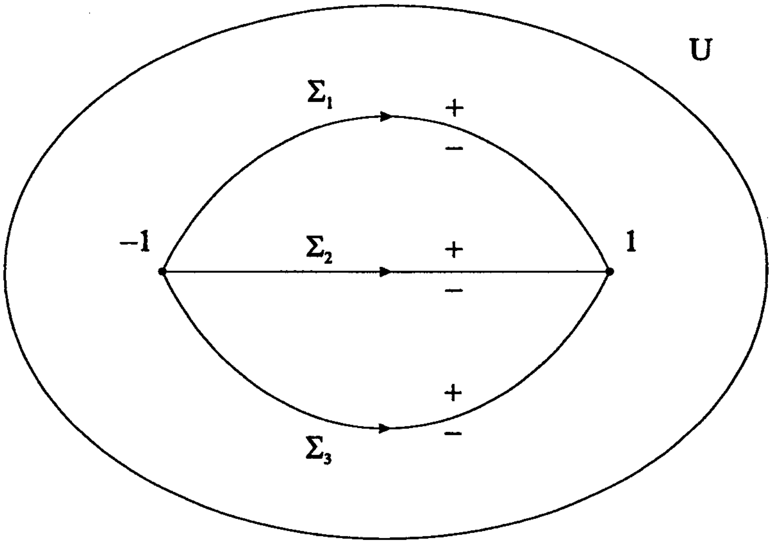


Fig. 3. Opening of lens in domain U.

(S-RH2) $S_+ = S_- \begin{pmatrix} 0 & w \\ -\frac{1}{w} & 0 \end{pmatrix}$ on $(-1,1)$.

$S_+ = S_- \begin{pmatrix} 1 & 0 \\ \frac{1}{w}\varphi^{-2n} & 1 \end{pmatrix}$ on Σ_1^o and Σ_3^o.

(S-RH3) $S(z) = I + \mathcal{O}\left(\frac{1}{z}\right)$ as $z \to \infty$.

(S-RH4) Conditions as $z \to 1$:

- For $\alpha < 0$:
$$S(z) = \mathcal{O}\begin{pmatrix} 1 & |z-1|^{\alpha} \\ 1 & |z-1|^{\alpha} \end{pmatrix}.$$
- For $\alpha = 0$:
$$S(z) = \mathcal{O}\begin{pmatrix} \log|z-1| & \log|z-1| \\ \log|z-1| & \log|z-1| \end{pmatrix}.$$
- For $\alpha > 0$:
$$S(z) = \begin{cases} \mathcal{O}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} & \text{as } z \to 1 \text{ outside the lens,} \\[2ex] \mathcal{O}\begin{pmatrix} |z-1|^{-\alpha} & 1 \\ |z-1|^{-\alpha} & 1 \end{pmatrix} & \text{as } z \to 1 \text{ inside the lens.} \end{cases}$$

(S-RH5) Similar conditions as $z \to -1$.

The endpoint condition (S-RH4) is rather awkward now, especially if $\alpha > 0$, where we distinguish between $z \to 1$ from within the lens, or from outside the lens. It turns out that they are necessary if we want a unique solution.

Exercise 16. Show that the Riemann-Hilbert problem (S-RH1)–(S-RH5) for S has a unique solution.
[Note: we already know that there is a solution, namely the one that we find after transformations $Y \mapsto T \mapsto S$. One way to prove that there is no other solution, is to show that these transformations are invertible. Another way is to assume that there is another solution $\tilde{S}$ and show that it must be equal to the S we already have.]

The opening of the lens in the transformation $T \mapsto S$ is also a crucial step in the papers [8, 9] by Deift et al., which deal with orthogonal polynomials on the real line, see also [7]. It transforms the oscillatory diagonal terms in the jump matrix for T into exponentially small off-diagonal terms in the jump matrix for S. Indeed, in (S-RH2) we have a jump matrix $\begin{pmatrix} 1 & 0 \\ \frac{1}{w}\varphi^{-2n} & 1 \end{pmatrix}$ on Σ_1^o and Σ_3^o. Since $|\varphi(z)| > 1$ for z on Σ_1^o and Σ_3^o, the entry $\frac{1}{w}\varphi^{-2n}$ in the jump matrix tends to 0 exponentially fast. The convergence is uniform on compact subsets of Σ_1^o and Σ_3^o, but it is not uniform near the endpoints ± 1.

9 Special case $\alpha = \beta = -\frac{1}{2}$

For special values of α and β, the subsequent analysis simplifies considerably. These are the cases $\alpha = \pm\frac{1}{2}$, $\beta = \pm\frac{1}{2}$. We will treat the case $\alpha = \beta = -\frac{1}{2}$, so that

$$w(z) = (1 - z^2)^{-\frac{1}{2}} h(z).$$

In this case, we open up the lens further so that Σ_1 and Σ_3 coincide along two intervals $[-1-\delta, -1]$ and $[1, 1+\delta]$, where $\delta > 0$ is some positive number.

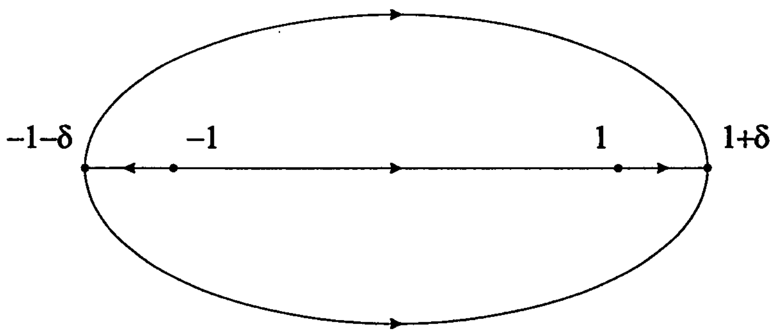


Fig. 4. Opening of lens in case $\alpha = \beta = -\frac{1}{2}$. The upper and lower lips of the lens coincide on the intervals $[-1-\delta, -1]$ and $[1, 1+\delta]$.

On the intervals $(-1-\delta, -1)$ and $(1, 1+\delta)$ two jumps are combined. If we calculate the total jump there, we have to be careful, since w has a jump

on these intervals too. In fact, we have

$$w_+(x) = -w_-(x), \qquad \text{for } x > 1 \text{ or } x < -1 \text{ (with } x \in U),$$

which follows from the fact that $\alpha = \beta = -\frac{1}{2}$. Then we calculate on $(-1-\delta, -1)$ or $(1, 1+\delta)$,

$$\begin{pmatrix} 1 & 0 \\ \frac{1}{w_-}\varphi^{-2n} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{1}{w_+}\varphi^{-2n} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \left(\frac{1}{w_-} + \frac{1}{w_+}\right)\varphi^{-2n} & 1 \end{pmatrix} = I.$$

This means that S is analytic across $(-1-\delta, -1)$ and $(1, 1+\delta)$. The only remaining jumps are on $[-1, 1]$ and on a simple closed contour that we call γ. We choose to orient γ in the positive direction (counterclockwise). It means that in the upper half-plane we have to reverse the orientation as shown in Figure 5.

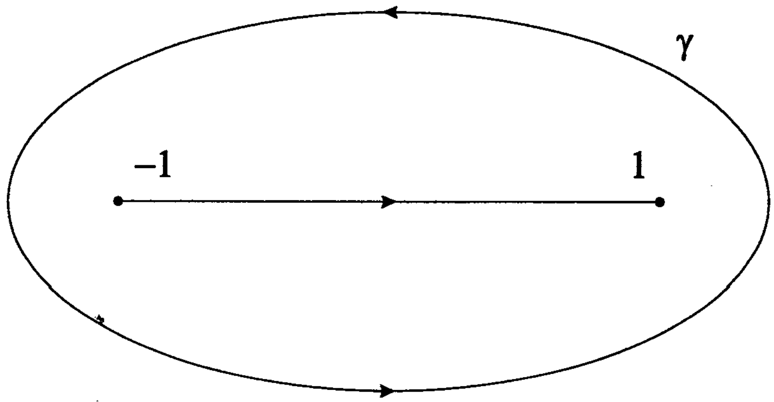


Fig. 5. Closed contour γ that encircles $[-1, 1]$. S has jumps only on γ and $[-1, 1]$.

It follows that in this special case S satisfies the following Riemann-Hilbert problem.

(S-RH1) S is analytic in $\mathbb{C} \setminus ([-1, 1] \cup \gamma)$.

(S-RH2) $S_+ = S_- \begin{pmatrix} 0 & w \\ -\frac{1}{w} & 0 \end{pmatrix}$ on $(-1, 1)$

$S_+ = S_- \begin{pmatrix} 1 & 0 \\ \frac{1}{w}\varphi^{-2n} & 1 \end{pmatrix}$ on $\gamma \cap \{\mathrm{Im} z < 0\}$ and

$S_+ = S_- \begin{pmatrix} 1 & 0 \\ -\frac{1}{w}\varphi^{-2n} & 1 \end{pmatrix}$ on $\gamma \cap \{\mathrm{Im} z > 0\}$.

(S-RH3) $S(z) = I + \mathcal{O}\left(\frac{1}{z}\right)$ as $z \to \infty$.

(S-RH4) $S(z) = \mathcal{O}\begin{pmatrix} 1 & |z-1|^{-\frac{1}{2}} \\ 1 & |z-1|^{-\frac{1}{2}} \end{pmatrix}$ as $z \to 1$.

(S-RH5) $S(z) = \mathcal{O}\begin{pmatrix} 1 & |z+1|^{-\frac{1}{2}} \\ 1 & |z+1|^{-\frac{1}{2}} \end{pmatrix}$ as $z \to -1$.

Exercise 17. If you do the analysis in this section for the case $\alpha = \beta = +\frac{1}{2}$ then everything will be the same except for the endpoint conditions (S-RH4) and (S-RH5). Show that they change to $S(z) = \mathcal{O}\begin{pmatrix} |z-1|^{-\frac{1}{2}} & 1 \\ |z-1|^{-\frac{1}{2}} & 1 \end{pmatrix}$ as $z \to 1$, and $S(z) = \mathcal{O}\begin{pmatrix} |z+1|^{-\frac{1}{2}} & 1 \\ |z+1|^{-\frac{1}{2}} & 1 \end{pmatrix}$ as $z \to -1$, respectively.

10 Model Riemann Hilbert problem

The jump matrix for S is uniformly close to the identity matrix on the simple closed contour γ. Only the jump on the interval $[-1,1]$ is not close to the identity. This suggests to look at the following model Riemann-Hilbert problem, where we ignore the jump on γ. We look for $N : \mathbb{C} \setminus [-1,1] \to \mathbb{C}^{2\times 2}$ satisfying

(N-RH1) N is analytic in $\mathbb{C} \setminus [-1,1]$.

(N-RH2) $N_+(x) = N_-(x) \begin{pmatrix} 0 & w(x) \\ -\frac{1}{w(x)} & 0 \end{pmatrix}$ for $x \in (-1,1)$.

(N-RH3) $N(z) \to I$ as $z \to \infty$.

(N-RH4) $N(z) = \mathcal{O}\begin{pmatrix} 1 & |z-1|^{-\frac{1}{2}} \\ 1 & |z-1|^{-\frac{1}{2}} \end{pmatrix}$ as $z \to 1$.

(N-RH5) $N(z) = \mathcal{O}\begin{pmatrix} 1 & |z+1|^{-\frac{1}{2}} \\ 1 & |z+1|^{-\frac{1}{2}} \end{pmatrix}$ as $z \to -1$.

The conditions (N-RH4) and (N-RH5) are specific for the weights under consideration (i.e., modified Jacobi weights with $\alpha = \beta = -\frac{1}{2}$). For more general weights on $[-1,1]$, the corresponding problem for N would include the parts (N-RH1), (N-RH2) and (N-RH3), but (N-RH4) and (N-RH5) have to be modified.

Exercise 18. Let N be given by

$$N(z) = \begin{pmatrix} \frac{a(z)+a^{-1}(z)}{2} & \frac{a(z)-a^{-1}(z)}{2i} \\ \frac{a(z)-a^{-1}(z)}{-2i} & \frac{a(z)+a^{-1}(z)}{2} \end{pmatrix},$$

where

$$a(z) = \frac{(z-1)^{1/4}}{(z+1)^{1/4}}.$$

Show that N satisfies parts (N-RH1), (N-RH2), and (N-RH3) with $w(x) \equiv 1$ (Legendre case). What would the conditions (N-RH4) and (N-RH5) be for this case?

The solution to the Riemann-Hilbert problem (N-RH1)–(N-RH5) is constructed with the use of the **Szegő function**. The Szegő function associated with a weight w on $[-1,1]$ is a scalar function $D : \mathbb{C} \setminus [-1,1] \to \mathbb{C}$ such that

(D-RH1) D is analytic and non-zero in $\mathbb{C} \setminus [-1,1]$.
(D-RH2) $D_+(x)D_-(x) = w(x)$ for $x \in (-1,1)$.
(D-RH3) the limit $\lim_{z\to\infty} D(z) = D_\infty$ exists and is a positive real number.

Note that (D-RH1)–(D-RH3) is a multiplicative scalar Riemann-Hilbert problem. We have not specified any endpoint conditions, so we cannot expect a unique solution. In general we want that $|D|$ behaves like $|w|^{1/2}$ also near the endpoints. So for a modified Jacobi weight we would add the endpoint conditions

(D-RH4) $D(z) = \mathcal{O}(|z-1|^{\alpha/2})$ as $z \to 1$,
(D-RH5) $D(z) = \mathcal{O}(|z+1|^{\beta/2})$ as $z \to -1$.

If the weight satisfies the Szegő condition

$$\int_{-1}^{1} \frac{\log w(x)}{\sqrt{1-x^2}} dx > -\infty$$

then the Szegő function exists and is given by

$$D(z) = \exp\left(\frac{(z^2-1)^{1/2}}{2\pi} \int_{-1}^{1} \frac{\log w(x)}{\sqrt{1-x^2}} \frac{dx}{x-z}\right). \tag{10.1}$$

Exercise 19. Show that $D(z)$ as given by (10.1) does indeed satisfy the jump condition $D_+D_- = w$.

Exercise 20. Show that the Szegő function for the pure Jacobi weight $w(x) = (1-x)^\alpha(1+x)^\beta$ is given by

$$D(z) = \left(\frac{(z-1)^\alpha(z+1)^\beta}{\varphi(z)^{\alpha+\beta}}\right)^{1/2},$$

with an appropriate branch of the square root.

Having D we can present the solution to the Riemann-Hilbert problem for N as follows.

$$N(z) = \begin{pmatrix} D_\infty & 0 \\ 0 & \frac{1}{D_\infty} \end{pmatrix} \begin{pmatrix} \frac{a(z)+a^{-1}(z)}{2} & \frac{a(z)-a^{-1}(z)}{2i} \\ \frac{a(z)-a^{-1}(z)}{-2i} & \frac{a(z)+a^{-1}(z)}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{D(z)} & 0 \\ 0 & D(z) \end{pmatrix}, \tag{10.2}$$

where

$$a(z) = \frac{(z-1)^{1/4}}{(z+1)^{1/4}}. \tag{10.3}$$

Exercise 21. Check that the jump condition (N-RH2) and endpoint conditions (N-RH4)–(N-RH5) are satisfied.
[Hint: The middle factor in the right-hand side of (10.2) appears as the solution for the Riemann-Hilbert problem for N in case $w \equiv 1$, see Exercise 18.]

Exercise 22. Show that $\det N(z) = 1$ for $z \in \mathbb{C} \setminus [-1,1]$.

11 Third transformation $S \mapsto R$

Now we can perform the final transformation $S \mapsto R$ in the case $\alpha = \beta = -\frac{1}{2}$. We define

$$R(z) = S(z) N^{-1}(z). \tag{11.1}$$

Since S and N have the same jump across $(-1,1)$ it is easy to see that $R_+(x) = R_-(x)$ for $x \in (-1,1)$, so that R is analytic across $(-1,1)$. Then R is analytic in $\mathbb{C} \setminus \gamma$ with possible singularities at ± 1. Since $\det N = 1$, we have from (N-RH4)

$$N^{-1}(z) = \mathcal{O}\begin{pmatrix} |z-1|^{-\frac{1}{2}} & |z-1|^{-\frac{1}{2}} \\ 1 & 1 \end{pmatrix}$$

as $z \to 1$. Thus

$$\begin{aligned} R(z) &= \mathcal{O}\begin{pmatrix} 1 & |z-1|^{-\frac{1}{2}} \\ 1 & |z-1|^{-\frac{1}{2}} \end{pmatrix} \mathcal{O}\begin{pmatrix} |z-1|^{-\frac{1}{2}} & |z-1|^{-\frac{1}{2}} \\ 1 & 1 \end{pmatrix} \\ &= \mathcal{O}\begin{pmatrix} |z-1|^{-\frac{1}{2}} & |z-1|^{-\frac{1}{2}} \\ |z-1|^{-\frac{1}{2}} & |z-1|^{-\frac{1}{2}} \end{pmatrix} \end{aligned}$$

as $z \to 1$. So all entries of R have an isolated singularity at $z = 1$ such that $R_{ij}(z) = \mathcal{O}\left(|z-1|^{-\frac{1}{2}}\right)$ as $z \to 1$. This implies that $z = 1$ is a removable singularity. Similarly it follows that $z = -1$ is a removable singularity.

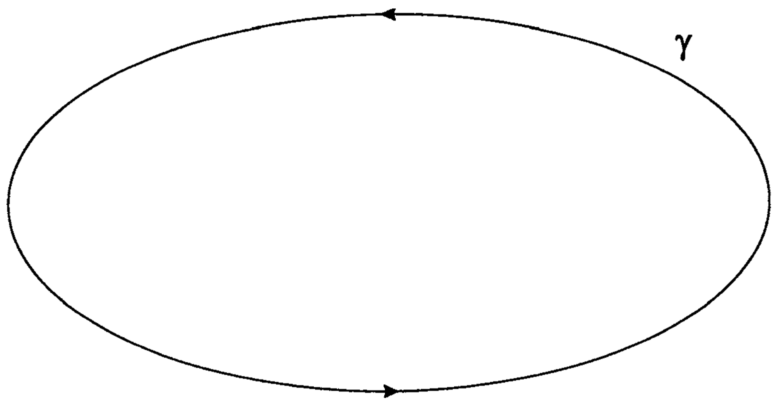


Fig. 6. Closed contour γ. R has a jump on γ only.

So R is analytic across the full interval $[-1,1]$, and so only has a jump on γ, as shown in Figure 6. We have the following Riemann-Hilbert problem for R.

(R-RH1) R is analytic on $\mathbb{C} \setminus \gamma$.
(R-RH2) $R_+(s) = R_-(s)V(s)$ where

$$V(s) = \begin{cases} N(s) \begin{pmatrix} 1 & 0 \\ \frac{1}{w(s)}\varphi(s)^{-2n} & 1 \end{pmatrix} N^{-1}(s) & \text{for } s \in \gamma \cap \{\operatorname{Im} z < 0\}, \\ N(s) \begin{pmatrix} 1 & 0 \\ -\frac{1}{w(s)}\varphi(s)^{-2n} & 1 \end{pmatrix} N^{-1}(s) & \text{for } s \in \gamma \cap \{\operatorname{Im} z > 0\}. \end{cases}$$

(R-RH3) $R(z) \to I$ as $z \to \infty$.

Observe that the jump matrix $V(s)$ is close to the identity matrix if n is large.

Exercise 23. Prove that V is analytic in a neighborhood Ω of γ, and that

$$\|V - I\|_\Omega = \mathcal{O}\left(e^{-cn}\right) \qquad \text{as } n \to \infty.$$

for some constant $c > 0$.

The Riemann-Hilbert problem for R is of the type discussed in Theorem 3.1. The problem is posed on a fixed contour γ (independent of n) and the jump matrix V is analytic in a neighborhood of γ where it is close to the identity. It follows from Theorem 3.1 that

$$R(z) = I + \mathcal{O}\left(e^{-cn}\right) \tag{11.2}$$

uniformly for z in $\mathbb{C} \setminus \gamma$. Tracing back the steps $Y \mapsto T \mapsto S \mapsto R$ we are then able to find asymptotics for Y as $n \to \infty$, and in particular for its $(1,1)$ entry, which is the orthogonal polynomial π_n.

Exercise 24. The analysis of Sections 9–11 goes through for all cases where the parameters α and β satisfy $\{\alpha, \beta\} \subset \{-\frac{1}{2}, \frac{1}{2}\}$. Work out the details for $\alpha = \beta = \frac{1}{2}$.
What goes wrong if $\alpha = \frac{3}{2}$?

12 Asymptotics for orthogonal polynomials (case $\alpha = \beta = -\frac{1}{2}$)

We repeat that the above analysis is valid for $\alpha = \beta = -\frac{1}{2}$, and according to the last exercise, can be extended to the cases $\alpha, \beta = \pm\frac{1}{2}$. Now we show how to get asymptotics from (11.2) for the orthogonal polynomials and for related quantities.

The easiest to obtain is asymptotics for $z \in \mathbb{C} \setminus [-1,1]$. For a given $z \in \mathbb{C} \setminus [-1,1]$, we can open the lens around $[-1,1]$ so that z is in the exterior of γ. Then by (7.1), (8.1), (11.1), and (11.2),

$$
\begin{aligned}
Y(z) &= \begin{pmatrix} 2^{-n} & 0 \\ 0 & 2^n \end{pmatrix} T(z) \begin{pmatrix} \varphi(z)^n & 0 \\ 0 & \varphi(z)^{-n} \end{pmatrix} \\
&= \begin{pmatrix} 2^{-n} & 0 \\ 0 & 2^n \end{pmatrix} S(z) \begin{pmatrix} \varphi(z)^n & 0 \\ 0 & \varphi(z)^{-n} \end{pmatrix} \\
&= \begin{pmatrix} 2^{-n} & 0 \\ 0 & 2^n \end{pmatrix} R(z) N(z) \begin{pmatrix} \varphi(z)^n & 0 \\ 0 & \varphi(z)^{-n} \end{pmatrix} \\
&= \begin{pmatrix} 2^{-n} & 0 \\ 0 & 2^n \end{pmatrix} \left(I + \mathcal{O}\left(e^{-cn}\right)\right) N(z) \begin{pmatrix} \varphi(z)^n & 0 \\ 0 & \varphi(z)^{-n} \end{pmatrix}
\end{aligned}
$$

as $n \to \infty$. For the orthogonal polynomial $\pi_n(z) = Y_{11}(z)$ we get

$$
\pi_n(z) = \left(\frac{\varphi(z)}{2}\right)^n \left[N_{11}(z)\left(1 + \mathcal{O}\left(e^{-cn}\right)\right) + N_{21}(z)\mathcal{O}\left(e^{-cn}\right)\right].
$$

Since N_{11} does not become zero in $\mathbb{C} \setminus [-1,1]$, we get the strong asymptotic formula

$$
\pi_n(z) = \left(\frac{\varphi(z)}{2}\right)^n N_{11}(z)\left(1 + \mathcal{O}\left(e^{-cn}\right)\right) \tag{12.1}
$$

as $n \to \infty$. For $N_{11}(z)$ we have from (10.2) the explicit expression

$$
N_{11}(z) = \frac{D_\infty}{D(z)} \frac{a(z) + a(z)^{-1}}{2} \tag{12.2}
$$

in terms of the Szegő function D associated with w and the function $a(z) = \frac{(z-1)^{1/4}}{(z+1)^{1/4}}$. The formula (12.1) is valid uniformly for z in compact subsets of $\overline{\mathbb{C}} \setminus [-1,1]$.

For asymptotics on the interval $[-1,1]$ we have to work somewhat harder, the basic difference being that the transformation from T to S is non-trivial now. We take z in the upper part of the lens. Then the transformations (7.1), (8.1), and (11.1) yield

$$\begin{aligned}
Y(z) &= \begin{pmatrix} 2^{-n} & 0 \\ 0 & 2^n \end{pmatrix} T(z) \begin{pmatrix} \varphi(z)^n & 0 \\ 0 & \varphi(z)^{-n} \end{pmatrix} \\
&= \begin{pmatrix} 2^{-n} & 0 \\ 0 & 2^n \end{pmatrix} S(z) \begin{pmatrix} 1 & 0 \\ \frac{1}{w(z)}\varphi(z)^{-2n} & 1 \end{pmatrix} \begin{pmatrix} \varphi(z)^n & 0 \\ 0 & \varphi(z)^{-n} \end{pmatrix} \\
&= \begin{pmatrix} 2^{-n} & 0 \\ 0 & 2^n \end{pmatrix} R(z)N(z) \begin{pmatrix} \varphi(z)^n & 0 \\ \frac{1}{w(z)}\varphi(z)^{-n} & \varphi(z)^{-n} \end{pmatrix} \\
&= \begin{pmatrix} 2^{-n} & 0 \\ 0 & 2^n \end{pmatrix} R(z) \begin{pmatrix} D_\infty & 0 \\ 0 & D_\infty^{-1} \end{pmatrix} \begin{pmatrix} \frac{a(z)+a^{-1}(z)}{2} & \frac{a(z)-a^{-1}(z)}{2i} \\ \frac{a(z)-a^{-1}(z)}{-2i} & \frac{a(z)+a^{-1}(z)}{2} \end{pmatrix} \\
&\qquad \times \begin{pmatrix} D(z)^{-1} & 0 \\ 0 & D(z) \end{pmatrix} \begin{pmatrix} \varphi(z)^n & 0 \\ \frac{1}{w(z)}\varphi(z)^{-n} & \varphi(z)^{-n} \end{pmatrix}.
\end{aligned}$$

So for the first column of Y we have

$$\begin{pmatrix} 2^n Y_{11}(z) \\ 2^{-n} Y_{21}(z) \end{pmatrix} = R(z) \begin{pmatrix} D_\infty & 0 \\ 0 & D_\infty^{-1} \end{pmatrix} \begin{pmatrix} \frac{a(z)+a^{-1}(z)}{2} & \frac{a(z)-a^{-1}(z)}{2i} \\ \frac{a(z)-a^{-1}(z)}{-2i} & \frac{a(z)+a^{-1}(z)}{2} \end{pmatrix} \begin{pmatrix} \frac{\varphi(z)^n}{D(z)} \\ \frac{D(z)}{w(z)\varphi(z)^n} \end{pmatrix}.$$

Now we take $x \in [-1,1]$, and we let z tend to x from the upper part of the lens. So we have to take the $+$boundary values of all quantities involved. It is tedious, but straightforward, to check that for $x \in (-1,1)$,

$$\varphi_+(x) = x + \sqrt{1-x^2} = \exp(i \arccos x),$$

$$\begin{aligned}
\frac{a_+(x) + a_+^{-1}(x)}{2} &= \frac{1}{\sqrt{2}(1-x^2)^{\frac{1}{4}}} \exp\left(\frac{1}{2} i \arccos x - i\frac{\pi}{4}\right), \\
\frac{a_+(x) - a_+^{-1}(x)}{2i} &= \frac{1}{\sqrt{2}(1-x^2)^{\frac{1}{4}}} \exp\left(-\frac{1}{2} i \arccos x + i\frac{\pi}{4}\right),
\end{aligned}$$

$$D_+(x) = \sqrt{w(x)} \exp(-i\psi(x)),$$

where $\psi(x)$ is a real-valued function, which is given by

$$\psi(x) = \frac{\sqrt{1-x^2}}{2\pi} \int_{-1}^{1} \frac{\log w(t)}{\sqrt{1-t^2}} \frac{dt}{t-x},$$

in which the integral is a principal value integral. Putting this all together we find for the orthogonal polynomial $\pi_n(x)$ with $x \in [-1,1]$,

$$\pi_n(x) = \frac{\sqrt{2}D_\infty}{2^n \sqrt{w(x)}(1-x^2)^{\frac{1}{4}}} \left[R_{11}(x) \cos\left(\left(n + \frac{1}{2}\right) \arccos x + \psi(x) - \frac{\pi}{4} \right) \right.$$
$$\left. - \frac{i}{D_\infty^2} R_{12}(x) \cos\left(\left(n - \frac{1}{2}\right) \arccos x + \psi(x) - \frac{\pi}{4} \right) \right], \quad (12.3)$$

where

$$R_{11}(x) = 1 + \mathcal{O}\left(e^{-cn}\right), \qquad R_{12}(x) = \mathcal{O}\left(e^{-cn}\right). \qquad (12.4)$$

The asymptotic formula (12.3)–(12.4) is valid uniformly for $x \in [-1,1]$. The fact that this includes the endpoints ± 1 is special to the case $\alpha = \beta = -\frac{1}{2}$. For more general α and β, the formula (12.3) continues to hold on compact subsets of the open interval $(-1,1)$, but with error terms $R_{11}(x) = 1 + \mathcal{O}(\frac{1}{n})$ and $R_{12}(x) = \mathcal{O}(\frac{1}{n})$. Near the endpoints ± 1, there is a different asymptotic formula.

The formula (12.3) clearly displays the oscillatory behavior of $\pi_n(x)$ on the interval $[-1,1]$. The amplitude of the oscillations is $\frac{\sqrt{2}D_\infty}{2^n \sqrt{w(x)}(1-x^2)^{\frac{1}{4}}}$ and it is easy to check that this remains bounded as $x \to \pm 1$. The main oscillating term is $\cos\left(\left(n + \frac{1}{2}\right) + \psi(x) - \frac{\pi}{4}\right)$ with corrections that are exponentially small as $n \to \infty$.

Exercise 25. The orthogonal polynomials for the weight $(1-x^2)^{-\frac{1}{2}}$ are the Chebyshev polynomials of the first kind $T_n(x)$ with the property

$$T_n(x) = \cos(n \arccos x), \qquad x \in [-1,1].$$

The monic Chebyshev polynomials are

$$\pi_n(x) = \frac{1}{2^{n-1}} T_n(x) \qquad \text{if } n \geq 1.$$

Compare this with the asymptotic formula (12.3). What are R_{11} and R_{12} in this case?
[Hint: It may be shown that for a Jacobi weight $(1-x)^\alpha (1+x)^\beta$ one has $D_\infty = 2^{-\frac{\alpha+\beta}{2}}$ and $\psi(x) = \frac{\alpha+\beta}{2} \arccos x - \frac{\alpha\pi}{2}$.]

Exercise 26. The formula (12.3)–(12.4) is also valid for the case $\alpha = \beta = \frac{1}{2}$. This may seem strange at first since then the amplitude of the oscillations $\frac{\sqrt{2}D_\infty}{2^n \sqrt{w(x)}(1-x^2)^{\frac{1}{4}}}$ is unbounded as $x \to \pm 1$. Still the formula (12.3) is valid uniformly on the closed interval $[-1,1]$. How can this be explained?

Exercise 27. Deduce from (12.1)–(12.4) that the coefficients a_n and b_n in the recurrence relation

$$x\pi_n(x) = \pi_{n+1}(x) + b_n\pi_n(x) + a_n^2\pi_{n-1}(x)$$

satisfy

$$a_n = \frac{1}{2} + \mathcal{O}\left(e^{-cn}\right), \qquad b_n = \mathcal{O}\left(e^{-cn}\right). \tag{12.5}$$

Remark related to exercise 27: Geronimo [19] made a thorough study of orthogonal polynomials with recurrence coefficients that approach their limits at an exponential rate. He showed that (12.5) holds, if and only if the underlying orthogonality measure is a modified Jacobi weight $(1-x)^{\pm\frac{1}{2}}(1+x)^{\pm\frac{1}{2}}h(x)$, plus at most a finite number of discrete masspoints outside $[-1,1]$. I thank Jeff Geronimo for this remark.

13 Case of general α and β

For the case of a modified Jacobi weight $(1-x)^\alpha(1+x)^\beta h(x)$ with general exponents $\alpha, \beta > -1$, we cannot do the transformation $T \mapsto S$ as described in Section 9. In general we have to stay with the transformation $T \mapsto S$ as in Section 8. So we are left with a Riemann-Hilbert problem on a system of contours as shown in Figure 3.

We continue to use the Szegő function $D(z)$ characterized by (D-RH1)–(D-RH5), and the solution of the model Riemann-Hilbert problem

$$N(z) = \begin{pmatrix} D_\infty & 0 \\ 0 & \frac{1}{D_\infty} \end{pmatrix} \begin{pmatrix} \frac{a(z)+a^{-1}(z)}{2} & \frac{a(z)-a^{-1}(z)}{2i} \\ \frac{a(z)-a^{-1}(z)}{-2i} & \frac{a(z)+a^{-1}(z)}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{D(z)} & 0 \\ 0 & D(z) \end{pmatrix},$$

with $a(z) = \frac{(z-1)^{\frac{1}{4}}}{(z+1)^{\frac{1}{4}}}$. Note that N satisfies

(N-RH1) N is analytic in $\mathbb{C} \setminus [-1,1]$.

(N-RH2) $N_+(x) = N_-(x) \begin{pmatrix} 0 & w(x) \\ -\frac{1}{w(x)} & 0 \end{pmatrix}$ for $x \in (-1,1)$.

(N-RH3) $N(z) \to I$ as $z \to \infty$.

The aim is again to prove that S is close to N if n is large. However, the attempt to define $R = SN^{-1}$ and prove that $R \sim I$ does not work. The problem lies near the endpoints ± 1, as SN^{-1} is not bounded near ± 1.

The way out of this is a **local analysis** near the endpoints ± 1. We are going to construct a so-called **local parametrix** P in a disk $\{|z-1| < \delta\}$ centered at 1, where δ is a small, but fixed, positive number. The parametrix should satisfy the following local Riemann-Hilbert problem

(P-RH1) P is analytic in $\{|z-1| < \delta\} \setminus \Sigma$ and continuous in $\{|z-1| \le \delta\} \setminus \Sigma$.

(P-RH2) P has the same jumps as S on $\Sigma \cap \{|z-1| < \delta\}$.
(P-RH3) $P = \left(I + \mathcal{O}\left(\frac{1}{n}\right)\right) N$ as $n \to \infty$, uniformly on $|z-1| = \delta$.
(P-RH4) P has the same behavior as S near 1.

Instead of an asymptotic condition, we now have in (P-RH3) a **matching condition**.

Similarly, we need a parametrix $\tilde{P}$ near -1 which should satisfy

($\tilde{\mathbf{P}}$-RH1) $\tilde{P}$ is analytic in $\{|z+1| < \delta\} \setminus \Sigma$ and continuous in $\{|z+1| \leq \delta\} \setminus \Sigma$.
($\tilde{\mathbf{P}}$-RH2) $\tilde{P}$ has the same jumps as S on $\Sigma \cap \{|z+1| < \delta\}$.
($\tilde{\mathbf{P}}$-RH3) $\tilde{P} = \left(I + \mathcal{O}\left(\frac{1}{n}\right)\right) N$ as $n \to \infty$, uniformly on $|z+1| = \delta$.
($\tilde{\mathbf{P}}$-RH4) $\tilde{P}$ has the same behavior as S near -1.

The construction of a local parametrix is done in [7, 9] with the help of Airy functions. Here we will need Bessel functions of order α. In the next section, we will outline the construction of P. In the remaining part of this section we will discuss how the transformation $S \mapsto R$ will be, assuming that we can find P and $\tilde{P}$.

We define R by

$$R(z) = \begin{cases} S(z)N(z)^{-1} & \text{if } |z-1| > \delta \text{ and } |z+1| > \delta, \\ S(z)P(z)^{-1} & \text{if } |z-1| < \delta, \\ S(z)\tilde{P}(z)^{-1} & \text{if } |z+1| < \delta. \end{cases} \tag{13.1}$$

Then R is analytic outside the system of contours γ shown in Figure 7.

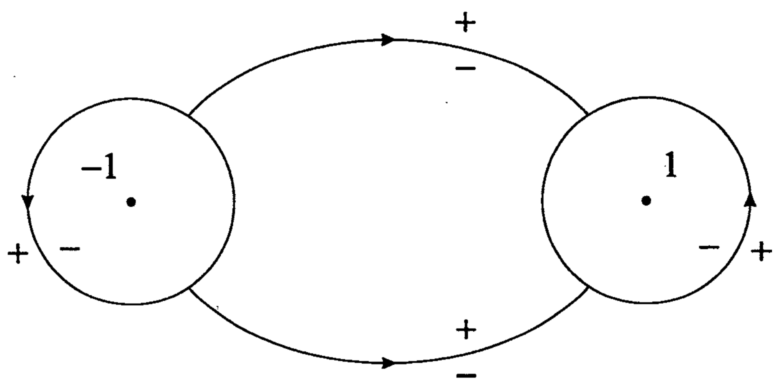


Fig. 7. System of contours γ so that R is analytic in $\mathbb{C} \setminus \gamma$. The system of contours γ consists of two circles of radius δ centered at ± 1, and two arcs joining these two circles.

R satisfies

(R-RH1) R is analytic on $\mathbb{C} \setminus \gamma$.
(R-RH2) $R_+ = R_- V$ on γ where

$$V = \begin{cases} PN^{-1} = I + \mathcal{O}\left(\frac{1}{n}\right) & \text{for } |z-1| = \delta, \\ \tilde{P}N^{-1} = I + \mathcal{O}\left(\frac{1}{n}\right) & \text{for } |z+1| = \delta, \\ N \begin{pmatrix} 1 & 0 \\ \frac{1}{w}\varphi^{-2n} & 1 \end{pmatrix} N^{-1} = I + \mathcal{O}\left(e^{-cn}\right) & \\ \qquad \text{on } (\Sigma_1 \cup \Sigma_2) \cap \{|z-1| > \delta, |z+1| > \delta\}. & \end{cases}$$

(R-RH3) $R(z) = I + \mathcal{O}\left(\frac{1}{z}\right)$ as $z \to \infty$.
(R-RH4) R remains bounded at the four points of self-intersection of γ.

Now the jump matrices are $I + \mathcal{O}\left(\frac{1}{n}\right)$ uniformly on γ. The contour γ is not a simple closed contour as in Theorem 3.1, so we cannot use that theorem directly. However, we can use the ideas in its proof to establish that we have $R(z) = I + \mathcal{O}\left(\frac{1}{n}\right)$.

Exercise 28. Prove that

$$R(z) = I + \mathcal{O}\left(\frac{1}{n}\right) \tag{13.2}$$

as $n \to \infty$, uniformly for $z \in \mathbb{C} \setminus \gamma$.

14 Construction of the local parametrix

The construction of the local parametrix P follows along a number of steps. More details can be found in [23].

Step 1: Reduction to constant jumps

We put for $z \in U \setminus (-\infty, 1]$,

$$W(z) = \left((z-1)^\alpha (z+1)^\beta h(z)\right)^{1/2},$$

where the branch of the square root is taken which is positive for $z > 1$. We seek P in the form

$$P = P^{(1)} \begin{pmatrix} W^{-1}\varphi^{-n} & 0 \\ 0 & W\varphi^n \end{pmatrix}.$$

In order to have (P-RH1), (P-RH2), and (P-RH4), we then get that $P^{(1)}$ should satisfy

(P1-RH1) $P^{(1)}$ is analytic in $\{|z-1| < \delta\} \setminus \Sigma$ and continuous in $\{|z-1| \leq \delta\} \setminus \Sigma$.

(P1-RH2) $P_+^{(1)} = P_-^{(1)} \begin{pmatrix} 1 & 0 \\ e^{\alpha\pi i} & 1 \end{pmatrix}$ on $\Sigma_1^o \cap \{|z-1| < \delta\}$,

$P_+^{(1)} = P_-^{(1)} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ on $(1-\delta, 1)$,

$P_+^{(1)} = P_-^{(1)} \begin{pmatrix} 1 & 0 \\ e^{-\alpha\pi i} & 1 \end{pmatrix}$ on $\Sigma_3^o \cap \{|z-1| < \delta\}$.

(P1-RH4) Conditions as $z \to 1$:

- If $\alpha < 0$, then $P^{(1)}(z) = \mathcal{O}\begin{pmatrix} |z-1|^{\alpha/2} & |z-1|^{\alpha/2} \\ |z-1|^{\alpha/2} & |z-1|^{\alpha/2} \end{pmatrix}$.
- If $\alpha = 0$, then $P^{(1)}(z) = \mathcal{O}\begin{pmatrix} \log|z-1| & \log|z-1| \\ \log|z-1| & \log|z-1| \end{pmatrix}$.
- If $\alpha > 0$, then

$$P^{(1)}(z) = \begin{cases} \mathcal{O}\begin{pmatrix} |z-1|^{\alpha/2} & |z-1|^{-\alpha/2} \\ |z-1|^{\alpha/2} & |z-1|^{-\alpha/2} \end{pmatrix} & \text{as } z \to 1 \text{ outside the lens,} \\ \\ \mathcal{O}\begin{pmatrix} |z-1|^{-\alpha/2} & |z-1|^{-\alpha/2} \\ |z-1|^{-\alpha/2} & |z-1|^{-\alpha/2} \end{pmatrix} & \text{as } z \to 1 \text{ inside the lens.} \end{cases}$$

For the moment we ignore the matching condition.

Step 2: Model Riemann-Hilbert problem

The constant jump problem we have for $P^{(1)}$ leads to a model problem for Ψ, defined in an auxiliary ζ-plane. The problem is posed on three semi-infinite rays γ_1, γ_2, and γ_3, where γ_2 is the negative real axis, $\gamma_1 = \{\arg\zeta = \sigma\}$, and $\gamma_3 = \{\arg\zeta = -\sigma\}$. Here σ is some angle in $(0, \pi)$, see Figure 8.

The Riemann-Hilbert problem for Ψ is:

(Ψ-RH1) Ψ is analytic in $\mathbb{C} \setminus (\gamma_1 \cup \gamma_2 \cup \gamma_3)$,

(Ψ-RH2) $\Psi_+ = \Psi_- \begin{pmatrix} 1 & 0 \\ e^{\alpha\pi i} & 1 \end{pmatrix}$ on γ_1^o,

$\Psi_+ = \Psi_- \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ on γ_2^o,

$\Psi_+ = \Psi_- \begin{pmatrix} 1 & 0 \\ e^{-\alpha\pi i} & 1 \end{pmatrix}$ on γ_3^o.

(Ψ-RH4) Conditions as $\zeta \to 0$:

- If $\alpha < 0$, then $\Psi(\zeta) = \mathcal{O}\begin{pmatrix} |\zeta|^{\alpha/2} & |\zeta|^{\alpha/2} \\ |\zeta|^{\alpha/2} & |\zeta|^{\alpha/2} \end{pmatrix}$.
- If $\alpha = 0$, then $\Psi(\zeta) = \mathcal{O}\begin{pmatrix} \log|\zeta| & \log|\zeta| \\ \log|\zeta| & \log|\zeta| \end{pmatrix}$.
- If $\alpha > 0$, then

$$\Psi(\zeta) = \begin{cases} \mathcal{O}\begin{pmatrix} |\zeta|^{\alpha/2} & |\zeta|^{-\alpha/2} \\ |\zeta|^{\alpha/2} & |\zeta|^{-\alpha/2} \end{pmatrix} & \text{as } \zeta \to 0 \text{ with } |\arg\zeta| < \sigma, \\ \mathcal{O}\begin{pmatrix} |\zeta|^{-\alpha/2} & |\zeta|^{-\alpha/2} \\ |\zeta|^{-\alpha/2} & |\zeta|^{-\alpha/2} \end{pmatrix} & \text{as } \zeta \to 0 \text{ with } \sigma < |\arg\zeta| < \pi. \end{cases}$$

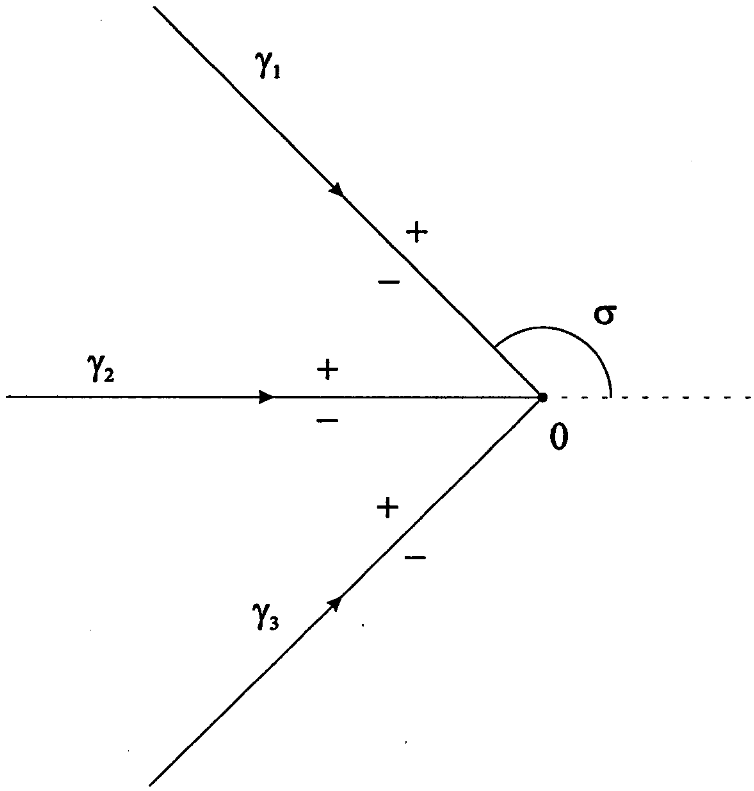


Fig. 8. Contours for the Riemann-Hilbert problem for Ψ.

There is no asymptotic condition (Ψ-RH3) for Ψ, so we cannot expect to have a unique solution. Indeed, there are in fact many solutions. In the next step we will construct one solution out of modified Bessel functions.

Step 3: Solution of model Riemann-Hilbert problem

The solution of the Riemann-Hilbert problem for Ψ will be built out of modified Bessel functions of order α, namely I_α and K_α. These are solutions of the modified Bessel differential equation

$$y'' + \frac{1}{\zeta} y' - \left(1 - \frac{\alpha^2}{\zeta^2}\right) y = 0.$$

The two functions $I_\alpha(2\zeta^{1/2})$ and $K_\alpha(2\zeta^{1/2})$ satisfy

$$y'' - \frac{1}{\zeta}\left(1 + \frac{\alpha^2}{4\zeta}\right) y = 0.$$

We consider these functions for $|\arg \zeta| < \pi$. On the negative real axis there is a jump. In fact we have the connection formulas, see [1, 9.6.30-31],

$$\begin{aligned}
I_\alpha(2\zeta^{1/2})_+ &= e^{\alpha\pi i} I_\alpha(2\zeta^{1/2})_- \\
K_\alpha(2\zeta^{1/2})_+ &= e^{-\alpha\pi i} K_\alpha(2\zeta^{1/2})_- - \pi i I_\alpha(2\zeta^{1/2})_-
\end{aligned}$$

for ζ on the negative real axis, oriented from left to right. We can put this in matrix-vector form

$$\left(I_\alpha(2\zeta^{1/2}) \ \tfrac{i}{\pi}K_\alpha(2\zeta^{1/2})\right)_+ = \left(I_\alpha(2\zeta^{1/2}) \ \tfrac{i}{\pi}K_\alpha(2\zeta^{1/2})\right)_- \begin{pmatrix} e^{\alpha\pi i} & 1 \\ 0 & e^{-\alpha\pi i} \end{pmatrix}.$$

Since the jump matrix is constant, it follows that the vector of derivatives satisfies the same jumps, and also if we multiply this vector by $2\pi i\zeta$. This has the effect of creating a matrix with determinant 1, due to the Wronskian relation [1, 9.6.15]). Thus

$$\begin{aligned}
&\begin{pmatrix} I_\alpha(2\zeta^{1/2}) & \frac{i}{\pi}K_\alpha(2\zeta^{1/2}) \\ 2\pi i\zeta^{1/2} I'_\alpha(2\zeta^{1/2}) & -2\zeta^{1/2}K'_\alpha(2\zeta^{1/2}) \end{pmatrix}_+ \\
&\quad = \begin{pmatrix} I_\alpha(2\zeta^{1/2}) & \frac{i}{\pi}K_\alpha(2\zeta^{1/2}) \\ 2\pi i\zeta^{1/2} I'_\alpha(2\zeta^{1/2}) & -2\zeta^{1/2}K'_\alpha(2\zeta^{1/2}) \end{pmatrix}_- \begin{pmatrix} e^{\alpha\pi i} & 1 \\ 0 & e^{-\alpha\pi i} \end{pmatrix}.
\end{aligned}$$

Now we have, as is easy to check,

$$\begin{pmatrix} e^{\alpha\pi i} & 1 \\ 0 & e^{-\alpha\pi i} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ e^{-\alpha\pi i} & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ e^{\alpha\pi i} & 1 \end{pmatrix}.$$

This last product consists exactly of the three jump matrices in the Riemann-Hilbert problem for Ψ. It follows that if we define Ψ by

$$\Psi(\zeta) = \begin{cases} \Psi_0(\zeta) & \text{for } |\arg\zeta| < \sigma, \\ \Psi_0(\zeta)\begin{pmatrix} 1 & 0 \\ -e^{\alpha\pi i} & 1 \end{pmatrix} & \text{for } \sigma < \arg\zeta < \pi, \\ \Psi_0(\zeta)\begin{pmatrix} 1 & 0 \\ e^{-\alpha\pi i} & 1 \end{pmatrix} & \text{for } -\pi < \arg\zeta < -\sigma, \end{cases} \tag{14.1}$$

where

$$\Psi_0(\zeta) = \begin{pmatrix} I_\alpha(2\zeta^{1/2}) & \frac{i}{\pi}K_\alpha(2\zeta^{1/2}) \\ 2\pi i\zeta^{1/2} I'_\alpha(2\zeta^{1/2}) & -2\zeta^{1/2}K'_\alpha(2\zeta^{1/2}) \end{pmatrix}, \tag{14.2}$$

then Ψ satisfies the jump condition (Ψ-RH2). Clearly, (Ψ-RH1) is also satisfied. Because of the known behavior of the modified Bessel functions near 0, see [1, 9.6.7–9], Ψ also has the behavior (Ψ-RH4) near 0.

Step 4: Construction of $P^{(1)}$

Define for $z \in \mathbb{C} \setminus (-\infty, 1]$,

$$f(z) = \frac{1}{4} \left[\log \varphi(z)\right]^2, \tag{14.3}$$

where we choose the principal branch of the logarithm. Since $\varphi_+(x)\varphi_-(x) = 1$ for $x \in (-1,1)$, we easily get that $f_+(x) = f_-(x)$. So f is analytic in $\mathbb{C} \setminus (-\infty, -1]$. The behavior near $z = 1$ is

$$f(z) = \frac{1}{2}(z-1) - \frac{1}{12}(z-1)^2 + \mathcal{O}((z-1)^3) \qquad \text{as } z \to 1.$$

So f is a conformal mapping of a neighborhood of 1 onto a neighborhood of 0. We choose $\delta > 0$ sufficiently small so that $\zeta = f(z)$ maps the disk $\{|z-1| < \delta\}$ conformally onto a convex neighborhood of 0 in the ζ-plane. We still have some freedom in the precise location of the contours Σ_1 and Σ_3. Here we use this freedom to specify that $\Sigma_1 \cap \{|z-1| < \delta\}$ should be mapped by f to a part of the ray $\arg \zeta = \sigma$ (we choose any $\sigma \in (0, \pi)$), and $\Sigma_3 \cap \{|z-1| < \delta\}$ to a part of the ray $\arg \zeta = -\sigma$.

Then $\Psi(n^2 f(z))$ satisfies the properties (P1-RH1), (P1-RH2), and (P1-RH4) of the Riemann-Hilbert problem for $P^{(1)}$. This would actually be the case for any choice of conformal map $\zeta = f(z)$, mapping $z = 1$ to $\zeta = 0$, and which is real and positive for $z > 1$. The specific choice of f is dictated by the matching condition for P, which we will look at in a minute. This will also explain the factor n^2. But this will not be enough to be able to do the matching. There is an additional freedom we have in multiplying $\Psi(n^2 f(z))$ **on the left** by an analytic factor. So we put

$$P^{(1)}(z) = E(z)\Psi(n^2 f(z)), \tag{14.4}$$

where E is an analytic 2×2 matrix valued function in $\{|z-1| < \delta\}$. It will depend on n. The precise form of E will be given in the next subsection.

Exercise 29. Show that for any analytic factor E the definition (14.4) gives a matrix valued function $P^{(1)}$ that satisfies the jump condition (P1-RH2) and the condition (P1-RH4) near 1.

Step 5: The matching condition

The parametrix P we now have is

$$P(z) = E(z)\Psi(n^2 f(z)) \begin{pmatrix} W(z)^{-1}\varphi(z)^{-n} & 0 \\ 0 & W(z)\varphi(z)^n \end{pmatrix} \tag{14.5}$$

where we have not specified E yet. The conditions (P-RH1), (P-RH2), and (P-RH4) are satisfied. We also have to take care of the matching condition

$$P(z) = \left(I + \mathcal{O}\left(\frac{1}{n}\right)\right) N(z) \qquad \text{for } |z-1| = \delta.$$

To achieve the matching, $E(z)$ should be close to

$$N(z) \begin{pmatrix} W(z)\varphi(z)^n & 0 \\ 0 & W(z)^{-1}\varphi(z)^{-n} \end{pmatrix} \left[\Psi(n^2 f(z))\right]^{-1}.$$

The idea is to replace Ψ here with an approximation Ψ^a. For fixed z with $|z-1| = \delta$, the function $\Psi(\zeta)$ is evaluated at $\zeta = n^2 f(z)$, which grows as $n \to \infty$. So to figure out what approximation Ψ^a to use, we need large ζ asymptotics of the modified Bessel functions and their derivatives. These functions have a known asymptotic expansion, see [1, 9.7.1–4]. From this it follows that

$$\Psi(\zeta) = \begin{pmatrix} \frac{1}{\sqrt{2\pi}}\zeta^{-1/4} & 0 \\ 0 & \sqrt{2\pi}\zeta^{1/4} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 + \mathcal{O}(\zeta^{-\frac{1}{2}}) & i + \mathcal{O}(\zeta^{-\frac{1}{2}}) \\ i + \mathcal{O}(\zeta^{-\frac{1}{2}}) & 1 + \mathcal{O}(\zeta^{-\frac{1}{2}}) \end{pmatrix} \begin{pmatrix} e^{2\zeta^{1/2}} & 0 \\ 0 & e^{-2\zeta^{1/2}} \end{pmatrix}.$$

Now we ignore the $\mathcal{O}(\zeta^{-\frac{1}{2}})$ terms, and we put

$$\Psi^a(\zeta) = \begin{pmatrix} \frac{1}{\sqrt{2\pi}}\zeta^{-1/4} & 0 \\ 0 & \sqrt{2\pi}\zeta^{1/4} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \begin{pmatrix} e^{2\zeta^{1/2}} & 0 \\ 0 & e^{-2\zeta^{1/2}} \end{pmatrix},$$

and then define

$$E(z) = N(z) \begin{pmatrix} W(z)\varphi(z)^n & 0 \\ 0 & W(z)^{-1}\varphi(z)^{-n} \end{pmatrix} \left[\Psi^a(n^2 f(z))\right]^{-1}.$$

Note that $e^{-2\zeta^{1/2}} = \varphi(z)^n$ for $\zeta = n^2 f(z)$. Thus

$$E(z) = N(z) \begin{pmatrix} W(z) & 0 \\ 0 & W(z)^{-1} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2\pi n} f(z)^{1/4} & 0 \\ 0 & \frac{1}{\sqrt{2\pi n}} f(z)^{-1/4} \end{pmatrix}. \quad (14.6)$$

The fact that the exponential factor $\varphi(z)^n$ gets cancelled is the reason for the choice of the mapping f and the factor n^2 in $\Psi(n^2 f(z))$. With this choice for E, it is easy to check that P satisfies the matching condition (P-RH3). We leave it as an exercise to show that E is analytic in a full neighborhood of 1. This completes the construction of the parametrix P in the neighborhood of 1.

A similar construction with modified Bessel functions of order β yields a parametrix $\tilde{P}$ in the neighborhood of -1.

Exercise 30.

(a) Show that $E_+(x) = E_-(x)$ for $x \in (1-\delta, 1)$, so that E is analytic across $(1-\delta, 1)$.
[Hint: On $(1-\delta, 1)$ we have $(f^{1/4})_+ = i(f^{1/4})_-$, $W_+W_- = w$, and $N_+ = N_- \begin{pmatrix} 0 & w \\ -\frac{1}{w} & 0 \end{pmatrix}$.]

(b) Show that the isolated singularity of E at 1 is removable.
[Hint: Use that $W(z)/D(z)$ is bounded and bounded away from zero near $z = 1$.]

15 Asymptotics for orthogonal polynomials (general case)

Knowing that we can construct the local parametrices P and $\tilde{P}$ we can go back to Section 13 and conclude that $R(z) = I + \mathcal{O}(\frac{1}{n})$ uniformly for $z \in \mathbb{C} \setminus \gamma$, where γ is the system of contours shown in Figure 7.

Then we can go back to our transformations $Y \mapsto T \mapsto S \mapsto R$, to obtain asymptotics for Y, and in particular for the orthogonal polynomial $\pi_n(z) = Y_{11}(z)$. We summarize here the results. For $z \in \mathbb{C} \setminus [-1, 1]$, we obtain

$$\pi_n(z) = \frac{\varphi(z)^n}{2^n} \frac{D_\infty}{D(z)} \frac{a(z) + a(z)^{-1}}{2} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right). \tag{15.1}$$

The $\mathcal{O}(\frac{1}{n})$ term is uniform for z in compact subsets of $\mathbb{C} \setminus [-1, 1]$. The formula is the same as the one (15.1) we found for the case $\alpha = \beta = -\frac{1}{2}$, except for the error term.

For $x \in (-1+\delta, 1-\delta)$, we obtain

$$\pi_n(x) = \frac{\sqrt{2} D_\infty}{2^n \sqrt{w(x)} (1-x^2)^{1/4}} \left(\cos\left(\left(n + \frac{1}{2}\right) \arccos x + \psi(x) - \frac{\pi}{4}\right) + \mathcal{O}\left(\frac{1}{n}\right)\right), \tag{15.2}$$

where $\psi(x) = -\arg D_+(x)$, compare also with (12.1).

Exercise 31. Check that we obtain (15.2) from taking the sum of $+$ and $-$ boundary values for the asymptotics (15.1) valid in $\mathbb{C} \setminus [-1, 1]$.

Near the endpoints ± 1 the asymptotic formula for $\pi_n(x)$ involves Bessel functions. For z in the upper part of the lens, inside the disk $\{|z-1| < \delta\}$, the expression for $Y(z)$ involves a product of no less than thirteen matrices (even after some simplifications). To summarize we have by (7.1), (8.1), and (13.1),

$$Y(z) = \begin{pmatrix} 2^{-n} & 0 \\ 0 & 2^n \end{pmatrix} R(z)P(z) \begin{pmatrix} 1 & 0 \\ \frac{1}{w}\varphi(z)^{-2n} & 1 \end{pmatrix} \begin{pmatrix} \varphi(z)^n & 0 \\ 0 & \varphi(z)^{-n} \end{pmatrix}, \tag{15.3}$$

with (due to (14.5), (14.6), (10.2), and (14.1)),

$$P(z) = E(z)\Psi(n^2 f(z)) \begin{pmatrix} W(z)^{-1}\varphi(z)^{-n} & 0 \\ 0 & W(z)\varphi(z)^n \end{pmatrix},$$

$$E(z) = N(z) \begin{pmatrix} W(z) & 0 \\ 0 & W(z)^{-1} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2\pi n} f(z)^{1/4} & 0 \\ 0 & \frac{1}{\sqrt{2\pi n}} f(z)^{-1/4} \end{pmatrix},$$

$$N(z) = \begin{pmatrix} D_\infty & 0 \\ 0 & D_\infty^{-1} \end{pmatrix} \begin{pmatrix} \frac{a(z)+a^{-1}(z)}{2} & \frac{a(z)-a^{-1}(z)}{2i} \\ \frac{a(z)-a^{-1}(z)}{-2i} & \frac{a(z)+a^{-1}(z)}{2} \end{pmatrix} \begin{pmatrix} D(z)^{-1} & 0 \\ 0 & D(z) \end{pmatrix},$$

and

$$\Psi(\zeta) = \begin{pmatrix} I_\alpha(2\zeta^{1/2}) & \frac{i}{\pi} K_\alpha(2\zeta^{1/2}) \\ 2\pi i \zeta^{1/2} I'_\alpha(2\zeta^{1/2}) & -2\zeta^{1/2} K'_\alpha(2\zeta^{1/2}) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -e^{\alpha\pi i} & 1 \end{pmatrix}.$$

We start to evaluate the product (15.3) at the right. Plugging in the formula for $P(z)$, we get

$$Y(z) = \begin{pmatrix} 2^{-n} & 0 \\ 0 & 2^n \end{pmatrix} R(z)E(z)\Psi(n^2 f(z)) \begin{pmatrix} W(z)^{-1} & 0 \\ \frac{W(z)}{w(z)} & W(z) \end{pmatrix}.$$

Since $W(z) = w(z)^{1/2} e^{\frac{1}{2}\alpha\pi i}$ in the region under consideration, we have

$$Y(z) = \begin{pmatrix} 2^{-n} & 0 \\ 0 & 2^n \end{pmatrix} R(z)E(z)\Psi(n^2 f(z)) \begin{pmatrix} e^{-\frac{1}{2}\alpha\pi i} & 0 \\ e^{\frac{1}{2}\alpha\pi i} & e^{\frac{1}{2}\alpha\pi i} \end{pmatrix} \begin{pmatrix} w(z)^{-1/2} & 0 \\ 0 & w(z)^{1/2} \end{pmatrix}. \tag{15.4}$$

Using the expression for $\Psi(\zeta)$, we get with $\zeta = n^2 f(z)$,

$$\begin{aligned} &\begin{pmatrix} 2^n & 0 \\ 0 & 2^{-n} \end{pmatrix} Y(z) \begin{pmatrix} w(z)^{1/2} & 0 \\ 0 & w(z)^{-1/2} \end{pmatrix} \\ &= R(z)E(z) \begin{pmatrix} I_\alpha(2\zeta^{1/2}) & \frac{i}{\pi} K_\alpha(2\zeta^{1/2}) \\ 2\pi i \zeta^{1/2} I'_\alpha(2\zeta^{1/2}) & -2\zeta^{1/2} K'_\alpha(2\zeta^{1/2}) \end{pmatrix} \begin{pmatrix} e^{-\frac{1}{2}\alpha\pi i} & 0 \\ 0 & e^{\frac{1}{2}\alpha\pi i} \end{pmatrix}. \end{aligned}$$

At this point we see that the first column of $Y(z)$ can be expressed in terms of I_α and I'_α only. It will not involve K_α and K'_α. Continuing now only with the first column and focusing on the $(1,1)$ entry, we get

$$\begin{pmatrix} \pi_n(z) \\ * \end{pmatrix} = \frac{1}{2^n w(z)^{1/2} e^{\frac{1}{2}\alpha\pi i}} R(z)E(z) \begin{pmatrix} 1 & 0 \\ 0 & 2\pi n f(z)^{1/2} \end{pmatrix} \begin{pmatrix} I_\alpha(2\zeta^{1/2}) \\ iI'_\alpha(2\zeta^{1/2}) \end{pmatrix}, \tag{15.5}$$

where $*$ denotes an unspecified entry. Looking now at the formula for $E(z)$, we see that we pick up an overall factor $\sqrt{2\pi n} f(z)^{1/4}$. We get from (15.5)

$$\begin{pmatrix} \pi_n(z) \\ * \end{pmatrix} = \frac{\sqrt{2\pi n} f(z)^{1/4}}{2^n w(z)^{1/2} e^{\frac{1}{2}\alpha\pi i}} R(z)N(z) \begin{pmatrix} W(z) & 0 \\ 0 & W(z)^{-1} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \begin{pmatrix} I_\alpha(2\zeta^{1/2}) \\ iI'_\alpha(2\zeta^{1/2}) \end{pmatrix}.$$

Next, we plug in the formula for N to obtain

$$\begin{pmatrix} \pi_n(z) \\ * \end{pmatrix} = \frac{\sqrt{2\pi n} f(z)^{1/4}}{2^n w(z)^{1/2} e^{\frac{1}{2}\alpha\pi i}} R(z) \begin{pmatrix} D_\infty & 0 \\ 0 & D_\infty^{-1} \end{pmatrix} \begin{pmatrix} \frac{a(z)+a^{-1}(z)}{2} & \frac{a(z)-a^{-1}(z)}{2i} \\ \frac{a(z)-a^{-1}(z)}{-2i} & \frac{a(z)+a^{-1}(z)}{2} \end{pmatrix}$$
$$\begin{pmatrix} \frac{W(z)}{D(z)} & 0 \\ 0 & \frac{D(z)}{W(z)} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \begin{pmatrix} I_\alpha(2\zeta^{1/2}) \\ iI'_\alpha(2\zeta^{1/2}) \end{pmatrix}, \tag{15.6}$$

where we still have $\zeta = n^2 f(z)$.

Now we choose $x \in (1-\delta, 1]$ and let $z \to x$ from within the upper part of the lens. The asymptotics for $\pi_n(x)$ will then involve the $+$boundary values of all functions appearing in (15.6). First we note that $f(x) = -\frac{1}{4}(\arccos x)^2$, so that

$$f(z)^{1/4} \to e^{\frac{1}{4}\pi i} \frac{\sqrt{\arccos x}}{\sqrt{2}}.$$

We also get that

$$\begin{pmatrix} I_\alpha(2\zeta^{1/2}) \\ iI'_\alpha(2\zeta^{1/2}) \end{pmatrix} \to e^{\frac{1}{2}\alpha\pi i} \begin{pmatrix} J_\alpha(n \arccos x) \\ J'_\alpha(n \arccos x) \end{pmatrix},$$

where J_α is the usual Bessel function. Note also that

$$\frac{a_+(x) + a_+^{-1}(x)}{2} = \frac{\exp\left(\frac{1}{2} i \arccos x - i\frac{\pi}{4}\right)}{\sqrt{2}(1-x^2)^{\frac{1}{4}}}$$

and

$$\frac{a_+(x) - a_+^{-1}(x)}{2i} = \frac{\exp\left(-\frac{1}{2} i \arccos x + i\frac{\pi}{4}\right)}{\sqrt{2}(1-x^2)^{\frac{1}{4}}}$$

so that

$$\begin{pmatrix} \frac{a(z)+a^{-1}(z)}{2} & \frac{a(z)-a^{-1}(z)}{2i} \\ \frac{a(z)-a^{-1}(z)}{-2i} & \frac{a(z)+a^{-1}(z)}{2} \end{pmatrix} \to \frac{1}{\sqrt{2}(1-x^2)^{\frac{1}{4}}} \begin{pmatrix} e^{\frac{1}{2} i \arccos x - i\frac{\pi}{4}} & e^{-\frac{1}{2} i \arccos x + i\frac{\pi}{4}} \\ -e^{-\frac{1}{2} i \arccos x + i\frac{\pi}{4}} & e^{\frac{1}{2} i \arccos x - i\frac{\pi}{4}} \end{pmatrix}.$$

Finally we have that $W_+(x) = \sqrt{w(x)}e^{\frac{1}{2}\alpha\pi i}$ and $D_+(x) = \sqrt{w(x)}e^{-i\psi(x)}$, so that

$$\frac{W(z)}{D(z)} \to e^{\frac{1}{2}\alpha\pi i+\psi(x)i}.$$

We get from (15.6)

$$\begin{aligned}
\begin{pmatrix}\pi_n(x)\\ *\end{pmatrix} &= \frac{\sqrt{\pi n\arccos x}\,e^{\frac{1}{4}\pi i}}{2^n\sqrt{w(x)}\sqrt{2}(1-x^2)^{1/4}}R(x)\begin{pmatrix}D_\infty & 0\\ 0 & D_\infty^{-1}\end{pmatrix}\\
&\quad\times\begin{pmatrix}e^{\frac{1}{2}i\arccos x-i\frac{\pi}{4}} & e^{-\frac{1}{2}i\arccos x+i\frac{\pi}{4}}\\ -e^{-\frac{1}{2}i\arccos x+i\frac{\pi}{4}} & e^{\frac{1}{2}i\arccos x-i\frac{\pi}{4}}\end{pmatrix}\\
&\quad\times\begin{pmatrix}e^{\frac{1}{2}\alpha\pi i+\psi(x)i} & 0\\ 0 & e^{-\frac{1}{2}\alpha\pi i-\psi(x)i}\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -i\\ -i & 1\end{pmatrix}\begin{pmatrix}J_\alpha(n\arccos x)\\ J_\alpha'(n\arccos x)\end{pmatrix}\\
&= \frac{\sqrt{\pi n\arccos x}}{2^n\sqrt{w(x)}(1-x^2)^{1/4}}R(x)\begin{pmatrix}D_\infty & 0\\ 0 & -iD_\infty^{-1}\end{pmatrix}\\
&\quad\times\begin{pmatrix}\cos(\zeta_1(x)) & \sin(\zeta_1(x))\\ \cos(\zeta_2(x)) & \sin(\zeta_2(x))\end{pmatrix}\begin{pmatrix}J_\alpha(n\arccos x)\\ J_\alpha'(n\arccos x)\end{pmatrix},
\end{aligned}\tag{15.7}$$

where

$$\zeta_1(x) = \frac{1}{2}\arccos x + \frac{1}{2}\alpha\pi + \psi(x), \qquad \zeta_2(x) = -\frac{1}{2}\arccos x + \frac{1}{2}\alpha\pi + \psi(x).$$

Exercise 32. Check that the formula for $\pi_n(x)$ remains bounded as $x \to 1$. [Hint: First note that $\frac{\sqrt{\arccos x}}{(1-x^2)^{1/4}}$ has a limit for $x \to 1$. Next, we should combine $\frac{1}{\sqrt{w(x)}}$ with the Bessel functions $J_\alpha(n\arccos x)$ and $J_\alpha'(n\arccos x)$. Since $J_\alpha(z) \sim \frac{1}{\Gamma(\alpha+1)}\left(\frac{z}{2}\right)^\alpha$ as $z \to 0$, we get that $\frac{J_\alpha(n\arccos x)}{\sqrt{w(x)}}$ has a limit as $x \to 1$. Finally, we should control $\frac{J_\alpha'(n\arccos x)}{\sqrt{w(x)}}$, which is unbounded as $x \to 1$ (unless $\alpha = 0$). However it gets multiplied by $\sin\zeta_1(x)$ and $\sin\zeta_2(x)$. It may be shown that $\zeta_j(x) = \mathcal{O}(\sqrt{1-x})$ as $x \to 1$ for $j = 1, 2$, and this is enough to show that $\sin\zeta_j(x)\frac{J_\alpha'(n\arccos x)}{\sqrt{w(x)}}$ remains bounded as well.]

Exercise 33. Show that, uniformly for θ in compact subsets of $\mathbb{C}$,

$$\lim_{n\to\infty}\frac{1}{n^\alpha 2^n}\binom{2n+\alpha+\beta}{n}\pi_n\left(\cos\frac{\theta}{n}\right) = C(h)\left(\frac{2}{\theta}\right)^\alpha J_\alpha(\theta), \tag{15.8}$$

where the constant $C(h)$ is given by

$$C(h) = \exp\left(\frac{1}{2\pi}\int_{-1}^{1}\frac{\log h(x) - \log h(1)}{\sqrt{1-x^2}}dx\right).$$

The limit (15.8) is the so-called Mehler-Heine formula, which is well-known for Jacobi polynomials, that is, for $h \equiv 1$, see, e.g., [1, 22.15.1] or [31]. From (15.8) one obtains the asymptotics of the largest zeros of π_n. Indeed, if $1 > x_1^{(n)} > x_2^{(n)} > \cdots$ denote the zeros of π_n, numbered in decreasing order, then (15.8) and Hurwitz's theorem imply that, for every $\nu \in \mathbb{N}$,

$$\lim_{n\to\infty} 2n^2 \left(1 - x_\nu^{(n)}\right) = j_{\alpha,\nu}^2,$$

where $0 < j_{\alpha,1} < j_{\alpha,2} < \cdots < j_{\alpha,\nu} < \cdots$ are the positive zeros of the Bessel function J_α. This property is well-known for Jacobi polynomials [1, 22.16.1].

Acknowledgements

I thank Kenneth McLaughlin, Walter Van Assche, and Maarten Vanlessen for allowing me to report on the material of [23] prior to its publication. I thank Maarten Vanlessen for his help in preparing the figures.

I am grateful to Erik Koelink and Walter Van Assche for organizing the Summer School in Orthogonal Polynomials and Special Functions and for giving me the opportunity to lecture there.

References

1. M. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, Dover Publications, New York, 1968.
2. A.I. Aptekarev, Sharp constants for rational approximations of analytic functions, Sbornik Math. **193** (2002), 3–72.
3. J. Baik, T. Kriecherbauer, K.T-R McLaughlin, and P.D. Miller, Uniform asymptotics for discrete orthogonal polynomials and applications to problems of probability, in preparation.
4. P. Bleher and A. Its, Semiclassical asymptotics of orthogonal polynomials, Riemann–Hilbert problem, and universality in the matrix model, Ann. Math. **150** (1999), 185–266.
5. P. Bleher and A. Its, Double scaling limit in the random matrix model: the Riemann-Hilbert approach, preprint math-ph/0201003.
6. K. Clancey and I. Gohberg, Factorization of matrix functions and singular integral operators Birkhäuser, Basel-Boston, 1981.
7. P. Deift, Orthogonal Polynomials and Random Matrices: A Riemann–Hilbert Approach, Courant Lecture Notes **3**, New York University, 1999.
8. P. Deift, T. Kriecherbauer, K.T-R McLaughlin, S. Venakides, and X. Zhou, Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory, Comm. Pure Appl. Math. **52** (1999), 1335–1425.
9. P. Deift, T. Kriecherbauer, K.T-R McLaughlin, S. Venakides, and X. Zhou, Strong asymptotics of orthogonal polynomials with respect to exponential weights, Comm. Pure Appl. Math. **52** (1999), 1491–1552.

10. P. Deift, T. Kriecherbauer, K.T-R McLaughlin, S. Venakides, and X. Zhou, A Riemann-Hilbert approach to asymptotic questions for orthogonal polynomials. J. Comput. Appl. Math. **133** (2001), 47–63.
11. P. Deift, S. Venakides, and X. Zhou, New results in small dispersion KdV by an extension of the steepest descent method for Riemann–Hilbert problems. Internat. Math. Res. Notices **1997**, no. 6, (1997), 286–299.
12. P. Deift and X. Zhou, A steepest descent method for oscillatory Riemann–Hilbert problems, Asymptotics for the MKdV equation, Ann. Math. **137** (1993), 295–368.
13. P. Deift and X. Zhou, Asymptotics for the Painlevé II equation, Comm. Pure Appl. Math. **48** (1995), 277–337.
14. P. Deift and X. Zhou, A priori L^p estimates for solutions of Riemann-Hilbert Problems, preprint, math.CA/0206224.
15. A.S. Fokas, A.R. Its, and A.V. Kitaev, The isomonodromy approach to matrix models in 2D quantum gravity, Comm. Math. Phys. **147** (1992), 395–430.
16. G. Freud, Orthogonale Polynome, Birkhäuser Verlag, Basel, 1969.
17. F. Gakhov, Boundary Value Problems, Pergamon Press, Oxford 1966. Reprinted by Dover Publications, New York, 1990.
18. W. Gawronski and W. Van Assche, Strong asymptotics for relativistic Hermite polynomials, Rocky Mountain J. Math, to appear.
19. J.S. Geronimo, Scattering theory, orthogonal polynomials, and q-series, SIAM J. Math. Anal. **25** (1994), 392-419.
20. T. Kriecherbauer and K.T-R McLaughlin, Strong asymptotics of polynomials orthogonal with respect to Freud weights. Internat. Math. Res. Notices **1999** (1999), 299–333.
21. A.B.J. Kuijlaars and K.T-R McLaughlin, Riemann–Hilbert analysis for Laguerre polynomials with large negative parameter, Comput. Meth. Funct. Theory **1** (2001), 205–233.
22. A.B.J. Kuijlaars and K.T-R McLaughlin, Asymptotic zero behavior of Laguerre polynomials with negative parameter, preprint math.CA/0205175.
23. A.B.J. Kuijlaars, K.T-R McLaughlin, W. Van Assche, and M. Vanlessen, The Riemann-Hilbert approach to strong asymptotics for orthogonal polynomials on [-1,1], preprint math.CA/0111252.
24. A.B.J. Kuijlaars and M. Vanlessen, Universality for eigenvalue correlations from the modified Jacobi unitary ensemble, Internat. Math. Res. Notices **2002**, no. 30, (2002), 1575–1600.
25. E. Levin and D.S. Lubinsky, Orthogonal Polynomials for Exponential Weights, CMS Books in Mathematics Vol. 4, Springer-Verlag, New York, 2001.
26. D.S. Lubinsky, Asymptotics of orthogonal polynomials: Some old, some new, some identities, Acta Appl. Math. **61** (2000), 207-256.
27. M.L. Mehta, Random Matrices, 2nd. ed. Academic Press, Boston, 1991.
28. N.I. Muskhelishvili, Singular Integral Equations, Noordhoff, Groningen, 1953. Reprinted by Dover Publications, New York, 1992.
29. P. Nevai, G. Freud, orthogonal polynomials and Christoffel functions. A case study, J. Approx. Theory **48** (1986), 3–167.
30. E.B. Saff and V. Totik, Logarithmic Potentials with External Fields, Springer-Verlag, Berlin, New York, 1997.
31. G. Szegő, Orthogonal Polynomials, Fourth edition, Colloquium Publications Vol. 23, Amer. Math. Soc., Providence R.I., 1975.

32. V. Totik, Weighted Approximation with Varying Weight, Lect. Notes Math. **1569**, Springer, Berlin, 1994.
33. W. Van Assche, Asymptotics for Orthogonal Polynomials, Lect. Notes. Math. **1265**, Springer-Verlag, Berlin, 1987.
34. W. Van Assche, J.S Geronimo, and A.B.J. Kuijlaars, Riemann–Hilbert problems for multiple orthogonal polynomials, pp. 23–59 in: "NATO ASI Special Functions 2000" (J. Bustoz et al. eds.), Kluwer Academic Publishers, Dordrecht 2001.

Exponential Asymptotics

Adri B. Olde Daalhuis

School of Mathematics, University of Edinburgh, JCMB, The King's Buildings, Edinburgh, EH9 3JZ, United Kingdom, adri@maths.ed.ac.uk

Summary. Recently, there has been a surge of practical and theoretical interest on the part of mathematical physicists, classical analysts and abstract analysts in the subject of exponential asymptotics, or hyperasymptotics, by which is meant asymptotic approximations in which the error terms are relatively exponentially small. Such approximations generally yield much greater accuracy than classical asymptotic expansions of Poincaré type, for which the error terms are algebraically small: in other words, they lead to "exponential improvement." They also enjoy greater regions of validity and yield a deeper understanding of other aspects of asymptotic analysis, including the Stokes phenomenon.

We shall obtain readily-applicable theories of hyperasymptotic expansions of solutions of differential equations and for integrals with one or more saddles. The main tool will be the Borel transform, which transforms the divergent asymptotic expansions into convergent series. Other methods will also be mentioned.

Topics to be discussed are:

- Least terms in divergent asymptotic expansions;
- Exponentially-improved asymptotic expansions;
- Smoothing of the Stokes phenomenon;
- Resurgence;
- Computation of Stokes multipliers (connection coefficients).

1 Introduction 212

1.1 Convergent Taylor series versus divergent asymptotic expansions . . 213
1.2 Notation and definitions 214
1.3 Divergent series and the Stokes phenomenon 215

2 Local hyperasymptotics for integrals 217

3 Global hyperasymptotics for integrals 221

4 Global hyperasymptotics for linear ODEs 226

4.1 Asymptotics for linear ODEs 226
4.2 Definitions and lemmas 227
4.3 The re-expansions 230
4.4 The optimal numbers of terms 233

E. Koelink and W. Van Assche (Eds.): LNM 1817, pp. 211–244, 2003.

4.5 The computation of the Stokes multipliers . . . 235
4.6 The computation of the hyperterminants . . . 236
4.7 Example . . . 237
5 Exercises . . . 241
References . . . 243

1 Introduction

In asymptotics one usually ignores terms that are relatively exponentially small. In many applications these exponentially small terms play an important role and that is why in the 1980's the field of exponential asymptotics was born. The exponentially small terms played an important role in a paper on scattering from a potential barrier (Pokrovskii and Khalatnikov), in work on adiabatic invariance (Meyer and Wasow), in extensive work on resonances (Simon, Harrell, Herbst and others), and in quantum chemistry (Connor and Brandas). Other applications are the growth of denditic crystals (Segur & Kruskal, McLeod & Amick and Levine) and viscous fingering (Tanveer). For more references see [12].

In exponentially improved asymptotic expansions, we include exponentially small terms via optimal truncation of the original divergent asymptotic expansion. One can re-expand the minimal remainder in a new divergent expansion. When we repeat this process of truncating and re-expanding, we obtain hyperasymptotic expansions.

There are two kind of hyperasymptotic expansions. In local hyperasymptotic expansions we use simple functions for the re-expansions. The drawback is that in general the coefficients in these re-expansions are not easy to compute, and the local hyperasymptotic expansions are valid only in small sectors. In chapter 2 we will show how we can obtain a local hyperasymptotic expansion for a simple integral.

The theory of global hyperasymptotics is introduced in chapter 3, where we obtain a hyperasymptotic expansion for an integral with saddles. The general theory of global hyperasymptotics is given in chapter 4, where we show that these expansions are valid in large sectors, and that the coefficients in each re-expansion are the same as or related to the coefficients used in the original Poincaré asymptotic expansion. This property is called *resurgence*. The other topics that we discuss in this chapter are the optimal number of terms at each level, and the computation of the so-called hyperterminants and Stokes multipliers.

We start this paper with an introduction into asymptotics. We first compare convergent Taylor series expansions with divergent asymptotic expansions. Then we give the notation and definitions that we will use in this paper and general theorems about asymptotic expansions. We finish the first chapter with a discussion of the Stokes phenomenon, which is the most important phe-

nomenon in asymptotics. It explains why we have to use nontrivial functions in global hyperasymptotics.

1.1 Convergent Taylor series versus divergent asymptotic expansions

$$\begin{aligned} F(x) &= \int_0^\infty e^{-xt}e^{-t}\,dt = \int_0^\infty e^{-xt}\bigl(1-\frac{t}{1!}+\frac{t^2}{2!}-\frac{t^3}{3!}+\cdots\bigr)\,dt \\ &= \frac{1}{x}-\frac{1}{x^2}+\frac{1}{x^3}-\frac{1}{x^4}+\cdots \qquad x\to\infty \qquad (1.1)\\ &= \sum_{s=0}^{n-1}(-)^s\frac{1}{x^{s+1}}+\varepsilon_n(x) \qquad (1.2)\end{aligned}$$

$$\begin{aligned} G(x) &= \int_0^\infty e^{-xt}\frac{1}{1+t}\,dt = \int_0^\infty e^{-xt}\bigl(1-t+t^2-t^3+\cdots\bigr)\,dt \\ &= \frac{1}{x}-\frac{1!}{x^2}+\frac{2!}{x^3}-\frac{3!}{x^4}+\cdots \qquad x\to\infty \qquad (1.3)\\ &= \sum_{s=0}^{n-1}(-)^s\frac{s!}{x^{s+1}}+R_n(x) \qquad (1.4)\end{aligned}$$

Differences

(i) Series (1.1) is convergent for $x>1$ and series (1.3) is divergent for all finite x.
(ii) In contrast to $\varepsilon_n(x)$ the remainder $R_n(x)$ cannot be expressed as the convergent sum of the tail.
(iii) When we take x fixed then $\varepsilon_n(x)$ diminishes steadily in size as n increases, whereas $R_n(x)$ diminishes steadily in size as long as n does not exceed $[x]$, thereafter $R_n(x)$ increases without limit. (See Figure 1.)

1.2 Notation and definitions

Let S denote the infinite sector $\alpha\le \mathrm{ph} z\le\beta$ and let $S(R)$ denote the annulus $|z|\ge R$, $z\in S$.

$$f(z)=\mathcal{O}(g(z)) \qquad \text{as} \quad |z|\to\infty \quad \text{in} \quad S, \qquad (1.5)$$

means that there exists a constant K such that $|f(z)|\le K|g(z)|$ as $|z|\to\infty$ in S.

$$f(z)=\mathcal{O}(g(z)) \qquad \text{when} \quad z\in S(R), \qquad (1.6)$$

means that there exists a constant K such that $|f(z)|\le K|g(z)|$ when $z\in S(R)$. We call the least upperbound of $|f(z)/g(z)|$ in $S(R)$ the *implied constant* of the $\mathcal{O}$-term for $S(R)$.

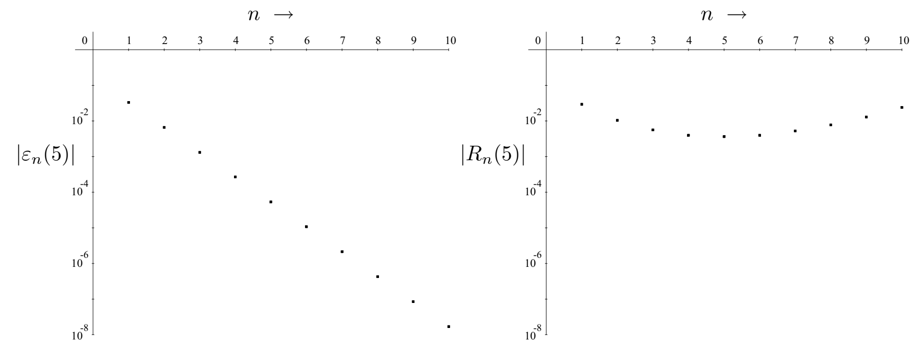


Fig. 1.

$$f(z) = o(g(z)) \qquad \text{as} \quad |z| \to \infty \quad \text{in} \quad S, \tag{1.7}$$

means that $f(z)/g(z) \to 0$ as $|z| \to \infty$ in S.

$$f(z) \sim g(z) \qquad \text{as} \quad |z| \to \infty \quad \text{in} \quad S, \tag{1.8}$$

means that $f(z)/g(z)$ tends to unity as $|z| \to \infty$ in S.

$$f(z) \sim \sum_{s=0}^{\infty} a_s z^{-s} \qquad \text{as} \quad |z| \to \infty \quad \text{in} \quad S, \tag{1.9}$$

means that for every fixed integer $n \geq 0$: $f(z) - \sum_{s=0}^{n-1} a_s z^{-s} = \mathcal{O}(z^{-n})$ as $|z| \to \infty$ in S. We call $\sum_{s=0}^{\infty} a_s z^{-s}$ an *asymptotic expansion* of $f(z)$, variable z the *asymptotic variable* (or *large parameter*) and the implied constant of the $\mathcal{O}$-term for an annulus $S(R)$ the nth *implied constant* of the asymptotic expansion for $S(R)$.

Theorem 1.1. (Uniqueness) *For a given function $f(z)$ in region S there is at most one expansion of the form* (1.9).

Theorem 1.2. *The nth implied constant of* (1.9) *for $S(R)$ cannot be less than $|a_n|$.*

Theorem 1.3. *If $f(z) \sim \sum_{s=0}^{\infty} a_s z^{-s}$ and $g(z) \sim \sum_{s=0}^{\infty} b_s z^{-s}$, as $|z| \to \infty$ in S, then*

$$\alpha f(z) + \beta g(z) \sim \sum_{s=0}^{\infty} (\alpha a_s + \beta b_s) z^{-s} \qquad \textit{as} \quad |z| \to \infty \quad \textit{in} \quad S,$$

providing α and β are constants; and

$$f(z)g(z) \sim \sum_{s=0}^{\infty} c_s z^{-s} \qquad \textit{as} \quad |z| \to \infty \quad \textit{in} \quad S,$$

where $c_s = \sum_{k=0}^{s} a_k b_{s-k}$.

Theorem 1.4. (Integration) *If* $f(z)$ *is continuous in* $S(R)$, *and if* $f(z) \sim \sum_{s=0}^{\infty} a_s z^{-s}$, *as* $|z| \to \infty$ *in* $S(R)$, *then*

$$\int_z^{\infty} \left[f(t) - a_0 - \frac{a_1}{t} \right] dt \sim \sum_{s=1}^{\infty} \frac{a_{s+1}}{s} z^{-s}, \qquad as \quad |z| \to \infty \quad in \quad S(R),$$

where the path of integration is the straight line joining z *to* ∞ *with a fixed argument.*

Theorem 1.5. (Differentiation 1) *Let* $f(x)$ *be continuously differentiable and* $f(x) \sim x^p$ *as* $x \to +\infty$, *where* $p \geq 1$. *Then* $f'(x) \sim px^{p-1}$, *provided that* $f'(x)$ *is nondecreasing for all sufficiently large* x.

Theorem 1.6. (Differentiation 2) *Assume that* $f(z) \sim \sum_{s=0}^{\infty} a_s z^{-s}$ *as* $|z| \to \infty$ *in* $S(R)$. *If* $f(z)$ *has a continuous derivative* $f'(z)$, *and if* $f'(z)$ *possesses an asymptotic expansion as* $|z| \to \infty$ *in* $S(R)$, *then*

$$f'(z) \sim -\sum_{s=1}^{\infty} s a_s z^{-s-1}, \qquad as \quad |z| \to \infty \quad in \quad S(R).$$

Theorem 1.7. (Differentiation 3) *Assume that* $f(z)$ *is analytic in* $S(R)$ *and that* $f(z) \sim \sum_{s=0}^{\infty} a_s z^{-s}$ *as* $|z| \to \infty$ *in* $S(R)$, *then*

$$f'(z) \sim -\sum_{s=1}^{\infty} s a_s z^{-s-1}, \qquad as \quad |z| \to \infty \quad in \quad S(R).$$

Theorem 1.8. *Let* $f(z)$ *be an analytic function in* $\tilde{S}(R) = \{z : |z| \geq R\}$, *and suppose that* $f(z) \sim \sum_{s=0}^{\infty} a_s z^{-s}$ *as* $|z| \to \infty$ *in* $\tilde{S}(R)$. *Then the asymptotic series is convergent and its sum is equal to* $f(z)$ *for all sufficiently large values of* z.

Theorem 1.9. *Let* $a_0, a_1, a_2, \cdots,$ *be an infinite sequence of arbitrary complex numbers, then there exists a function* $f(z)$ *such that* $f(z) \sim \sum_{s=0}^{\infty} a_s z^{-s}$ *as* $|z| \to \infty$ *in* S.

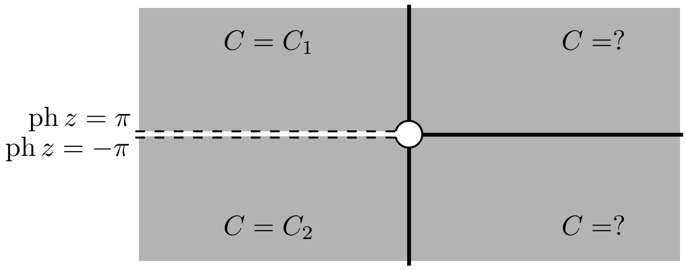
1.3 Divergent series and the Stokes phenomenon

The simple ordinary differential equation $w'(z) = \dfrac{e^z}{z}$ has formal series solutions

$$w(z, C) \sim e^z \sum_{n=0}^{\infty} \frac{n!}{z^{n+1}} \quad + \quad C, \qquad \text{as } |z| \to \infty.$$

The constant C is visible when $\mathrm{Re}z < 0$ and hidden by the divergent series when $\mathrm{Re}z > 0$. What happens with the 'constant' C when we travel from sector $\mathrm{ph}z \in (\frac{1}{2}\pi, \pi)$ to sector $\mathrm{ph}z \in (-\pi, -\frac{1}{2}\pi)$? Stokes noticed in [13] that

in general $C_1 \neq C_2$, and we will call this phenomenon the *Stokes phenomenon.*

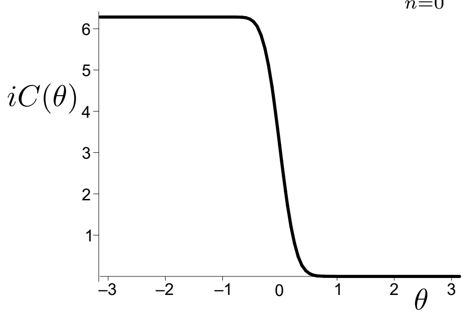


Stokes (1902): "The inferior term enters as it were into a mist, is hidden for a little from view, and comes out with its coefficient changed". Like Stokes in [13] we will make the constant C visible via optimal truncation. Fix $|z|$ for the moment. The optimal number of terms in the divergent series is $N = [|z|]$. 'Take' $C_1 = 0$. Then as $|z| \to \infty$

$$w(z,0) = e^z \sum_{n=0}^{N-1} \frac{n!}{z^{n+1}} \quad + \quad e^{z-|z|}\mathcal{O}\left(z^{-1/2}\right), \qquad\qquad 0 < \mathrm{ph} z \leq \pi.$$

Hence, in this sector the remainder is smaller than a constant, and C should be visible.

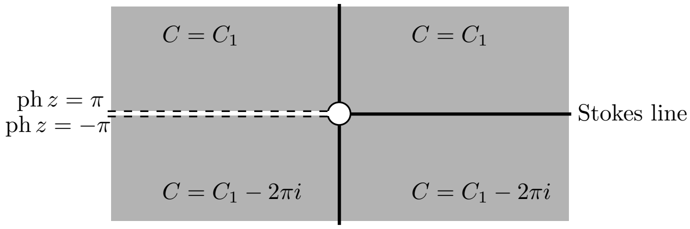
Write $z = |z|e^{i\theta}$, and plot $C(\theta) = w(z,0) - e^z \sum_{n=0}^{N-1} \frac{n!}{z^{n+1}}$.



Hence, the change in 'constant' C behaves like a so-called error function. Note that the change in C really happens at the line ph$z = 0$. We will call this line (direction) the *Stokes line*. Hence, the complete picture of the asymptotic expansions of the solutions of the simple ordinary differential equation is:

$$w(z, C_1) \sim e^z \sum_{n=0}^{\infty} \frac{n!}{z^{n+1}} \quad + \quad C,$$

where C is



As we have seen above C is a function of θ, the phase of z. Near the Stokes line $C(\theta)$ behaves like an error function. In 1989 Berry [2] introduced the *Smoothing of the Stokes Phenomenon* by showing that

$$C(\theta) = -\pi i \, \mathrm{erfc}(\tilde{\theta}),$$

where $\tilde{\theta}$ is a (complicated) function of θ, and where erfc(t) is the complementary error function. For a wide class of problems the smoothing of the Stokes phenomenon can be described via an error function.

The main observation that we should make is that the asymptotics near the Stokes lines is highly non-trivial. This will justify the non-trivial functions that we will use in hyperasymptotic expansions.

2 Local hyperasymptotics for integrals

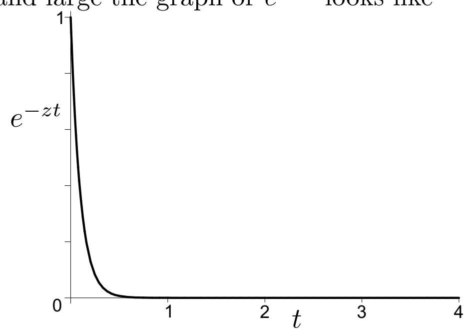
In this chapter we will construct a hyperasymptotic expansion for a simple example which is valid in a small sector. Let us consider the simple function

$$f(z) = \int_0^\infty \frac{e^{-zt}}{1+t} \, dt, \tag{2.1}$$

which can be written as an exponential integral, an incomplete gamma function, or as a confluent hypergeometric function:

$$f(z) = e^z E_1(z) = e^z \Gamma(0, z) = U(1, 1, z). \tag{2.2}$$

For z positive and large the graph of e^{-zt} looks like



Hence, the main contribution to the integral (2.1) comes from $t = 0$, and we will expand $\dfrac{1}{1+t} = 1 - t + t^2 - \cdots + (-t)^{N-1} + \dfrac{(-t)^N}{1+t}$ near $t = 0$. The

result is a truncated asymptotic expansion with an integral representation for the remainder:

$$f(z) = \sum_{n=0}^{N-1} \frac{(-)^n n!}{z^{n+1}} + R_0(z,N), \tag{2.3}$$

where

$$R_0(z,N) = (-)^N \int_0^\infty \frac{e^{-zt} t^N}{1+t}\, dt. \tag{2.4}$$

We obtain a sharp estimate for the remainder via

$$|R_0(z,N)| \le \int_0^\infty e^{-zt} t^N\, dt = \frac{N!}{z^{N+1}}. \tag{2.5}$$

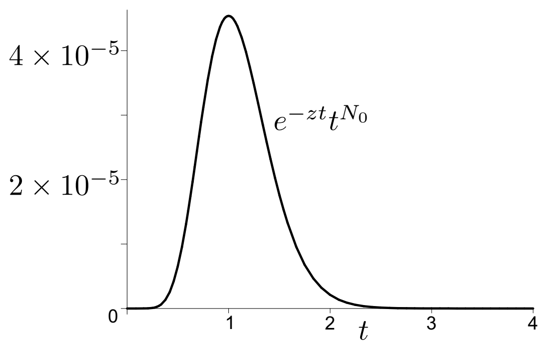
Hence, in this simple example the first omitted term in the asymptotic expansion is a 'sharp' upper bound for the remainder.

What is the best approximation that we can obtain by truncating this divergent asymptotic expansion? We regard the remainder as a function of N. It is easy to check that $N \mapsto \dfrac{N!}{z^{N+1}}$ has a minimum near $N \approx z$. Let us take $N_0 = [z]$, the integer part of z. With this choice we obtain the remainder estimate:

$$|R_0(z,N_0)| \le \frac{N_0!}{z^{N_0+1}} \approx \frac{\Gamma(z+1)}{z^{z+1}} \sim \sqrt{\frac{2\pi}{z}} e^{-z}, \qquad \text{as } z \to \infty. \tag{2.6}$$

The minimal remainder is exponentially small, hence, it is much smaller than any term in the divergent asymptotic expansion. Note that we obtained this remarkable result by allowing N to be a function of z.

Can we obtain better approximations? This is not possible via the simple divergent asymptotic expansion. We have to truncate this expansion, and re-expand the remainder. The dominating part in integral representation (2.4) is



Now the main contribution to the integral comes from $t = 1$. To obtain the re-expansion we expand

$$\frac{1}{1+t} = \frac{1}{2}\frac{1}{1-\frac{1-t}{2}} = \sum_{n=0}^{N-1} \frac{(1-t)^n}{2^{n+1}} + \frac{\left(\frac{1-t}{2}\right)^N}{1+t} \tag{2.7}$$

near $t = 1$. The result is

$$f(z) = \sum_{n=0}^{N_0-1} \frac{(-)^n n!}{z^{n+1}} + \sum_{n=0}^{N-1} 2^{-n-1} p_n(z, N_0) + R_1(z, N_0, N), \tag{2.8}$$

where

$$R_1(z, N_0, N) = \int_0^\infty \frac{e^{-zt}(-t)^{N_0}\left(\frac{1-t}{2}\right)^N}{1+t}\, dt, \tag{2.9}$$

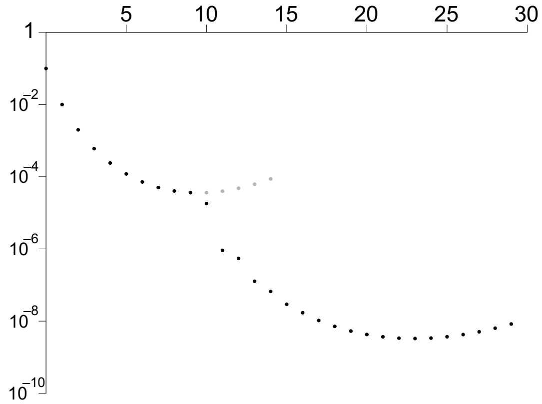
and

$$p_n(z, N_0) = \int_0^\infty e^{-zt} t^{N_0} (1-t)^n \, dt, \tag{2.10}$$

which can be seen as a finite sum of factorials, and computed via the recurrence relation

$$\begin{aligned} z p_{n+1}(z, N_0) &= (z - N_0 - n - 1) p_n(z, N_0) + n p_{n-1}(z, N_0), \\ p_0(z, N_0) &= \frac{N_0!}{z^{N_0+1}}, \qquad p_1(z, N_0) = \frac{N_0!}{z^{N_0+2}}(z - N_0 - 1). \end{aligned} \tag{2.11}$$

The absolute values of the terms in *level one hyperasymptotic expansion* (2.8) are



This figure clearly shows the improvement that we obtain via the re-expansion, and that the re-expansion is again divergent.

Let us again try to determine the optimal number of terms in the re-expansion. Recall that $N_0 \sim z$ and write $N = \rho z$. Then

$$|R_1(z, N_0, N)| \le \int_0^\infty e^{-zt} t^{N_0} \left| \frac{t-1}{2} \right|^N dt \approx \int_0^\infty e^{-zh(t)}\, dt, \tag{2.12}$$

where

$$h(t) = t - \ln t - \rho \ln \left| \frac{t-1}{2} \right|. \tag{2.13}$$

This function has two saddle points, and the dominating saddle point for the asymptotics of the final integral in (2.12) is the one in interval $t > 1$. This

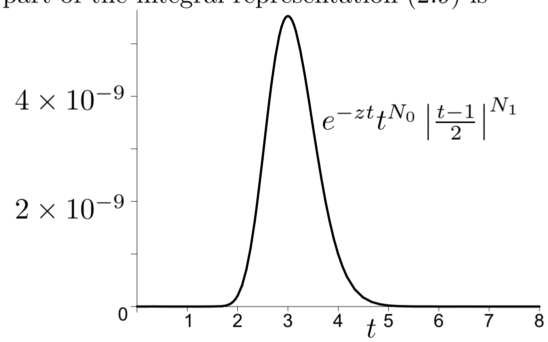
saddle point will depend on ρ, and it is not difficult to show that the optimal choice for $\rho = \frac{4}{3}$. In that case the two saddle points for $h(t)$ are located at $t = \frac{1}{3}$ and $t = 3$.

We take $N = N_1 = \left[\frac{4}{3}z\right]$. Applying the saddle point method to the integral in (2.12) with this choice of N gives us the result

$$|R_1(z, N_0, N_1)| \sim \frac{e^{-3z}3^z}{\sqrt{z}}\mathcal{O}(1) = \frac{e^{-1.901\cdots z}}{\sqrt{z}}\mathcal{O}(1), \qquad \text{as } z \to \infty. \tag{2.14}$$

Hence, we have almost doubled the precision of our approximation.

We can continue this process and obtain higher level re-expansions. The dominating part of the integral representation (2.9) is



Now the main contribution to the integral comes from $t = 3$. To obtain the next re-expansion we expand

$$\frac{1}{1+t} = \frac{1}{4}\frac{1}{1-\frac{3-t}{4}} = \sum_{n=0}^{N-1}\frac{(3-t)^n}{4^{n+1}} + \frac{\left(\frac{3-t}{4}\right)^N}{1+t} \tag{2.15}$$

near $t = 3$. The process is similar to the previous re-expansion, but we omit the details. The remainder in this second level hyperasymptotic expansion will be of the order

$$|R_2(z, N_0, N_1, N_2)| \sim \frac{e^{-3.59\cdots z}}{\sqrt{z}}\mathcal{O}(1), \qquad \text{as } z \to \infty. \tag{2.16}$$

This simple process can be continued and will give better and better approximations. The advantage of such a hyperasymptotic expansion is that it is in terms of very simple functions. The reason that we call it local hyperasymptotics is that it is valid only in 'small' z-sectors. The global hyperasymptotic expansions that we will discus in the next sections will have large z-sectors of validity, and these sectors will contain several Stokes lines. In [4] and [5] it is explained how to obtain these local hyperasymptotic expansions for confluent hypergeometric functions, and more general problems. In the general case the computations of the coefficients in the re-expansions might also be difficult. The global hyperasymptotic expansions will incorporate *resurgence*, which means that no new coefficients have to be computed in the re-expansions.

3 Global hyperasymptotics for integrals

In this chapter we will construct hyperasymptotic expansions for integrals with simple saddle points. Let

$$I(z) = \int_C e^{zh(t)} g(t)\,dt \tag{3.1}$$

be an integral in which $h(t)$ has only simple saddle points, that are, $t_1, \cdots, t_p$, such that $h'(t_j) = 0$ and $h''(t_j) \neq 0$. For simplicity we will assume that $h(t)$ and $g(t)$ are entire functions. We define

$$\lambda_n = h(t_n), \qquad \lambda_{nm} = \lambda_n - \lambda_m, \tag{3.2}$$

and

$$T^{(n)}(z) = \sqrt{z} e^{-\lambda_n z} I(z) = \sqrt{z} \int_{C_n(\theta)} e^{z(h(t)-\lambda_n)} g(t)\,dt, \tag{3.3}$$

where contour $C_n(\theta)$ is a steepest descent contour in the complex plane that passes through t_n, and in the case that there is more than one saddle point on this contour, t_n is the dominating one for the asymptotics of $T^{(n)}(z)$. Steepest descent paths will depend on $\theta = \mathrm{ph} z$. On the steepest descent paths we want $z(h(t) - \lambda_n) \leq 0$.

The example in this section will be

$$h(t) = \frac{1}{5}t^5 - \frac{3}{4}t^4 + t^3 - \frac{3}{2}t^2 + 2t, \tag{3.4}$$

with saddle points at

$$t_1 = 1, \quad t_2 = 2, \quad t_3 = i, \quad t_4 = -i, \tag{3.5}$$

and

$$\lambda_1 = \frac{19}{20}, \quad \lambda_2 = \frac{2}{5}, \quad \lambda_3 = \frac{3}{4} + \frac{6}{5}i, \quad \lambda_4 = \frac{3}{4} - \frac{6}{5}i, \tag{3.6}$$

see Figure 2.

Before we discus hyperasymptotics for these integrals we have to define *adjacent* saddles. Consider all the steepest descent paths through t_n for different θ. Some of these are special in that they encounter other saddle points t_m. We call these saddle points *adjacent to* t_n. The phase θ for which $t_m \in C_n(\theta)$ is $\theta_{nm} := -\mathrm{ph}(\lambda_{nm})$.

These steepest descent paths show that t_1 is adjacent to all the others, t_2 is only adjacent to t_1, t_3 is adjacent to t_1 and t_4, and t_4 is adjacent to t_1 and t_3.

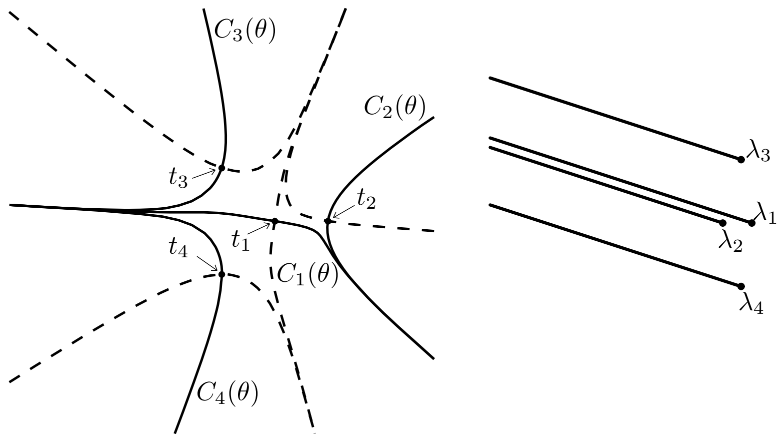


Fig. 2. The saddle points and steepest descent paths for the case $\theta = \frac{1}{10}\pi$ (left) and the $h(t)$-images of these steepest descent paths (right). The dashed lines are the steepest ascent paths.

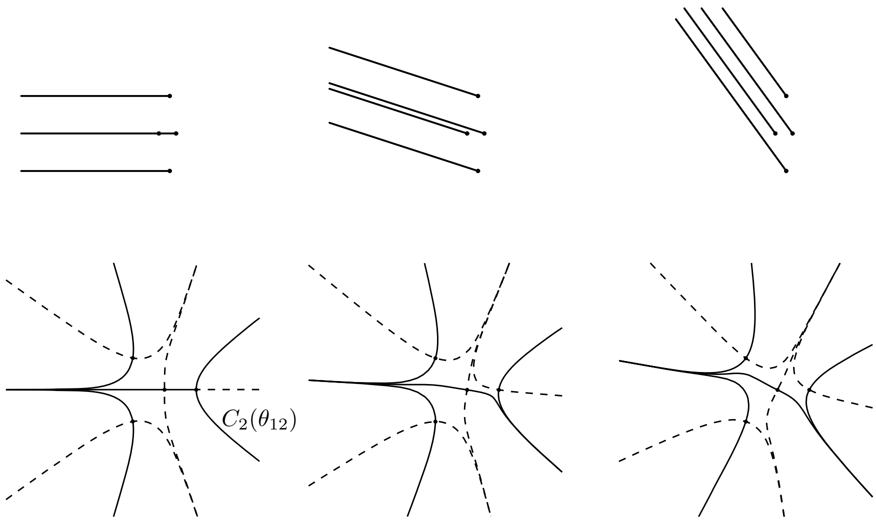


Fig. 3. The steepest descent paths and their $h(t)$-images for the cases $\theta = 0$ (left), $\theta = \frac{1}{10}\pi$ (middle), and $\theta = \frac{3}{10}\pi$ (right).

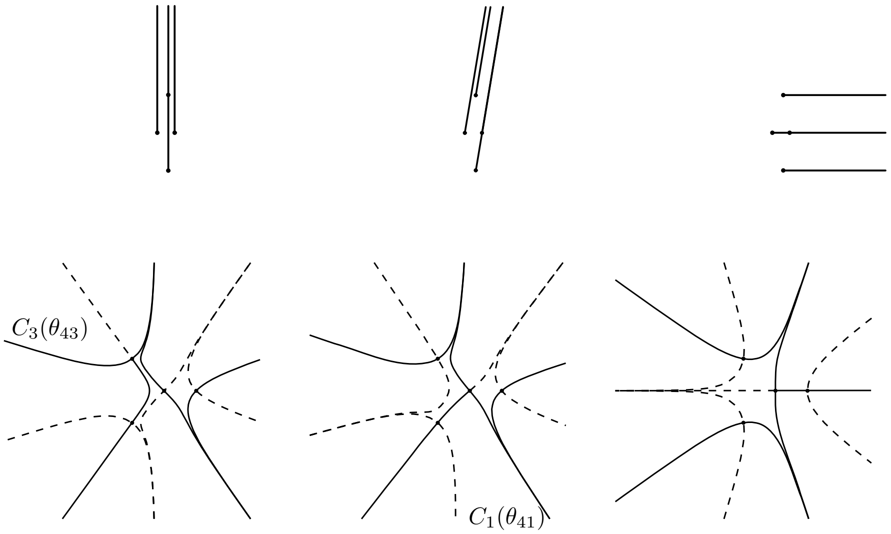


Fig. 4. The steepest descent paths and their $h(t)$-images for the cases $\theta = \frac{5}{10}\pi$ (left), $\theta = \arctan(6)$ (middle), and $\theta = \pi$ (right).

Remarks.

(1) Note that the contours in the cases $\theta = 0$ and $\theta = \pi$ are the same except that the steepest descent and steepest ascent contours are interchanged.

(2) The contours for the case $\theta = \frac{5}{10}\pi$ illustrate the Stokes phenomenon: Suppose that our integral is $T^{(4)}(z)$, with θ a bit smaller than $\frac{5}{10}\pi$. Hence, this integral is determined uniquely by the fact that $C_4(\theta)$ starts in the valley at $\infty e^{-7\pi i/10}$ and terminates in the valley at $\infty e^{-11\pi i/10}$. This doesn't change when θ becomes a bit bigger than $\frac{5}{10}\pi$. However, in that case the only way to travel from the starting valley to the terminating valley via steepest descent contours is first taken $C_4(\theta)$ in the clockwise direction and then $C_3(\theta)$ in the counterclockwise direction. Hence, the line $\theta = \frac{5}{10}\pi$ is a Stokes line, and crossing this line switches on the exponentially small term $-e^{\lambda_{34} z \pi i} T^{(3)}(z)$.

The construction of the hyperasymptotic expansion starts with the integral representation

$$T^{(n)}(z) = \frac{1}{2\pi i} \int_0^\infty \int_{\Gamma_n(\theta)} \frac{e^{-u} u^{-1/2} \left(\lambda_n - h(t)\right)^{-1/2} g(t)}{1 - \frac{u}{z(\lambda_n - h(t))}} \, dt \, du, \qquad (3.7)$$

where $\Gamma_n(\theta)$ is a contour that encircles $C_n(\theta)$. The proof of this integral representation is given in [3] and in the exercises. By expanding the denominator we obtain the truncated asymptotic expansion

$$T^{(n)}(z) = \sum_{s=0}^{N-1} \frac{a_{sn}}{z^s} + R_n^{(0)}(z,N), \tag{3.8}$$

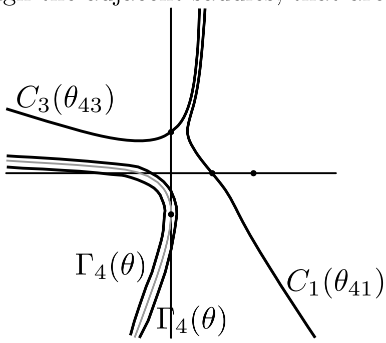
where the coefficients can be computed via the integral representation

$$a_{sn} = \frac{\Gamma(s+\frac{1}{2})}{2\pi i} \oint_{t_n} \frac{g(t)}{(\lambda_n - h(t))^{s+\frac{1}{2}}}\, dt, \tag{3.9}$$

where the contour of integration is a simple loop encircling t_n. The advantage of integral representation (3.7) is that we obtain (3.9) for free, and that we obtain a simple integral representation for the remainder:

$$R_n^{(0)}(z,N) = z^{-N} \frac{1}{2\pi i} \int_0^\infty \int_{\Gamma_n(\theta)} \frac{e^{-u} u^{N-1/2} (\lambda_n - h(t))^{-N-1/2} g(t)}{1 - \frac{u}{z(\lambda_n - h(t))}}\, dt\, du. \tag{3.10}$$

The next step is crucial. We deform $\Gamma_n(\theta)$ into the union of steepest descent contours through the adjacent saddles, that are, $C_m(\theta_{nm})$.



$$\begin{aligned} R_n^{(0)}(z,N) = &\sum_{\text{adjacent } m} \frac{(-)^{\gamma_{nm}}}{z^N} \frac{1}{2\pi i} \\ &\times \int_0^\infty \int_{C_m(\theta_{nm})} \frac{e^{-u} u^{N-1/2} (\lambda_n - h(t))^{-N-1/2} g(t)}{1 - \frac{u}{z(\lambda_n - h(t))}}\, dt\, du, \end{aligned} \tag{3.11}$$

where γ_{nm} is an 'orientation anomaly'. Finally, we use the substitution $u = t_0(\lambda_n - h(t))$ and obtain

$$R_n^{(0)}(z,N) = \sum_{\text{adjacent } m} \frac{(-)^{\gamma_{nm}} z^{1-N}}{2\pi i} \int_0^{\infty/\lambda_{nm}} \frac{e^{-\lambda_{nm} t_0} t_0^{N-1}}{z - t_0} T^{(m)}(t_0)\, dt_0, \tag{3.12}$$

where we have used integral representation (3.3) for the t-integral.

We summarise the above in

$$T^{(n)}(z) = \sum_{s=0}^{N-1} \frac{a_{sn}}{z^s} + \sum_{\text{adjacent } m} \frac{(-)^{\gamma_{nm}} z^{1-N}}{2\pi i} \int_0^{\infty/\lambda_{nm}} \frac{e^{-\lambda_{nm} t_0} t_0^{N-1}}{z - t_0} T^{(m)}(t_0)\, dt_0. \tag{3.13}$$

The only thing that we have to do to obtain a re-expansion is to substitute (3.13) with $n = m$ into the integrals on the right-hand side of (3.12). We obtain

$$R_n^{(0)}(z, N) = \sum_{\text{adjacent } m} \frac{(-)^{\gamma_{nm}} z^{1-N}}{2\pi i} \sum_{s=0}^{N_m - 1} a_{sm} \int_0^{\infty/\lambda_{nm}} \frac{e^{-\lambda_{nm} t_0} t_0^{N-s-1}}{z - t_0}\, dt_0 + R_n^{(1)}(z, N), \tag{3.14}$$

where

$$R_n^{(1)}(z, N) = \sum_{\substack{m \text{ adjacent} \\ \text{to } n}} \sum_{\substack{k \text{ adjacent} \\ \text{to } m}} \frac{(-)^{\gamma_{nm} + \gamma_{mk}} z^{1-N}}{(2\pi i)^2} \times \int_0^{\infty/\lambda_{nm}} \int_0^{\infty/\lambda_{mk}} \frac{e^{-\lambda_{nm} t_0 - \lambda_{mk} t_1} t_0^{N - N_m} t_1^{N_m - 1}}{(z - t_0)(t_0 - t_1)} T^{(k)}(t_1)\, dt_1\, dt_0. \tag{3.15}$$

Remarks.

(1) Re-expansion (3.14) is in terms of non trivial special functions. The integral in (3.14) is the so-called level one *hyperterminant*. It is the simplest function with a Stokes phenomenon.

$$\int_0^{\infty/\lambda} \frac{e^{-\lambda t} t^\alpha}{z - t}\, dt = -(-z)^\alpha e^{-z\lambda} \Gamma(1+\alpha) \Gamma(-\alpha, -z\lambda).$$

(2) Note that the coefficients in re-expansion (3.14) are the coefficients of the other integrals. Hence, no new coefficients have to be computed.

(3) In the exercises you will show that

$$a_{Nn} \sim - \sum_{\text{adjacent } m} \frac{(-)^{\gamma_{nm}}}{2\pi i} \sum_{s=0}^{\infty} \frac{a_{sm} \Gamma(N - s)}{\lambda_{nm}^{N-s}}, \qquad \text{as } N \to \infty. \tag{3.16}$$

Hence, in the asymptotic expansions of the late coefficients the early coefficients reappear. This is called *resurgence*.

In the next section we will discus the main details of global hyperasymptotic expansions. These are: the optimal number of terms at each level, the estimates for remainder terms, the computation of hyperterminants, and the computation of Stokes multipliers (not encountered in this section).

We finish this section by noting that we can obtain the next re-expansion by the substitution of (3.13) into the integral of (3.15). This process can be continued, and we would obtain the complete hyperasymptotic expansion.

4 Global hyperasymptotics for linear ODEs

4.1 Asymptotics for linear ODEs

We shall investigate solutions of differential equations of the form

$$\frac{d^n w}{dz^n} + f_{n-1}(z)\frac{d^{n-1}w}{dz^{n-1}} + \cdots + f_0(z)w = 0, \tag{4.1}$$

in which the coefficients $f_m(z)$, $m = 0, 1, \cdots, n-1$ can be expanded in power series

$$f_m(z) = \sum_{s=0}^{\infty} \frac{f_{sm}}{z^s}, \tag{4.2}$$

that converge on an open annulus $|z| > a$, and the point at infinity is an irregular singularity of rank 1. Formal series solutions in descending powers of z are given by

$$e^{\lambda_j z} z^{\mu_j} \sum_{s=0}^{\infty} a_{sj} z^{-s}, \qquad j = 1, 2, \cdots, n. \tag{4.3}$$

The constants λ_j, μ_j and a_{sj} are found by substituting into the differential equation and equating coefficients after setting $a_{0j} = 1$. In this way we obtain the characteristic equation

$$\sum_{m=0}^{n} \lambda_j^m f_{0m} = 0, \tag{4.4}$$

where we take $f_{0n} = 1$, to compute λ_j. The constants μ_j are given by

$$\mu_j = -\left(\sum_{m=0}^{n-1} \lambda_j^m f_{1m}\right) \Big/ \left(\sum_{m=1}^{n} m\lambda_j^{m-1} f_{0m}\right). \tag{4.5}$$

For the coefficients a_{sj} we obtain the recurrence relation

$$\begin{aligned}(s-1)a_{s-1,j} &\sum_{m=1}^{n} m\lambda_j^{m-1} f_{0,m} = \\ &\sum_{t=2}^{s} a_{s-t,j} \sum_{p=0}^{t} (\mu_j + t - s)_p \sum_{m=p}^{n} \binom{m}{p} \lambda_j^{m-p} f_{t-p,m},\end{aligned} \tag{4.6}$$

where Pochhammer's symbol $(\alpha)_p$ is defined by $(\alpha)_p = \Gamma(\alpha+p)/\Gamma(\alpha)$.

We shall impose the restriction

$$\lambda_j \neq \lambda_k, \qquad j \neq k. \tag{4.7}$$

This restriction ensures that the left-hand side of (4.6 does not vanish. We shall also assume that the μ_j are nonintegers. In [8] it is shown that this restriction is not needed for the final results.

4.2 Definitions and lemmas

We define

$$\left.\begin{aligned} \theta_{kj} &= \operatorname{ph}(\lambda_j - \lambda_k), \\ \lambda_{kj} &= \lambda_k - \lambda_j, \\ \mu_{kj} &= \mu_k - \mu_j, \end{aligned}\right\} \qquad j \neq k, \qquad \text{and} \quad \tilde{\mu} = \max\{\mathrm{Re}\mu_1, \cdots, \mathrm{Re}\mu_n\}. \tag{4.8}$$

We call

$$\eta \in \mathbb{R} \text{ is } \textit{admissible} \iff \eta \neq \theta_{kj} \ (\mathrm{mod}\ 2\pi), \qquad 1 \leq j,k \leq n,\ j \neq k, \tag{4.9}$$

see Figure 5.

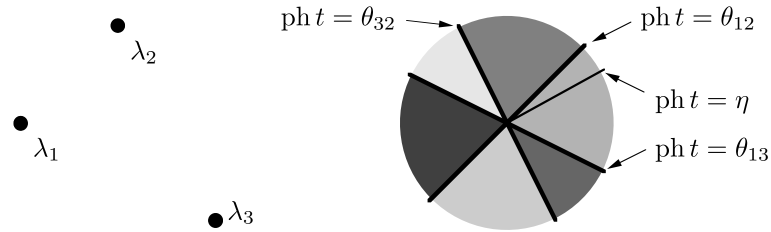


Fig. 5. An example of admissible and nonadmissible directions.

For fixed admissible η we consider a t-plane together with parallel cuts from each λ_k to ∞ along the ray $\operatorname{ph}(t-\lambda_k) = \eta$. See Figure 6. If we specify for $k = 1, 2, \cdots, n$,

$$\log(t - \lambda_k) = \log|t - \lambda_k| + i\eta, \tag{4.10}$$

for all t such that $\operatorname{ph}(t-\lambda_k) = \eta$, then we denote the t-plane with these cuts and choices of logarithms by $\mathcal{P}_\eta$. Thus $\log(t-\lambda_k)$ is continuous within $\mathcal{P}_\eta$, and is defined by (4.10) on $\operatorname{ph}(t-\lambda_k) = \eta$.

Let η be admissible. Then we define

$$\eta^- = \inf\{\check{\eta} < \eta \mid \tilde{\eta} \text{ is admissible for all } \tilde{\eta} \in (\check{\eta}, \eta]\}, \tag{4.11}$$

$$\eta^+ = \sup\{\hat{\eta} > \eta \mid \tilde{\eta} \text{ is admissible for all } \tilde{\eta} \in [\eta, \hat{\eta})\}, \tag{4.12}$$

$$\mathcal{I}_\eta = (\eta^-, \eta^+). \tag{4.13}$$

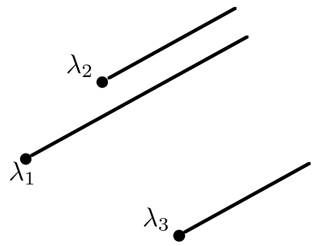


Fig. 6. Cuts for $\mathcal{P}_\eta$.

Note that $\eta^\pm$ are not admissible. In Figure 5 η^- is the value $\theta_{13} \mod 2\pi$ for which $|\eta-\theta_{13}|$ is least, and η^+ is the value $\theta_{12} \mod 2\pi$ for which $|\theta_{12}-\eta|$ is least.

With these definitions for $\eta^\pm$ we define the z-sectors

$$\mathcal{S}(\eta)=\{z\,|\,\mathrm{Re}(ze^{i\eta})<-a \text{ and } \frac{\pi}{2}-\eta^+<\mathrm{ph}z<\frac{3\pi}{2}-\eta^-\}, \tag{4.14}$$

$$\overline{\mathcal{S}}(\eta)=\{z\,|\,\mathrm{Re}(ze^{i\eta})<-a \text{ and } \pi-\eta^+\le\mathrm{ph}z\le\pi-\eta^-\}, \tag{4.15}$$

The main tools that we will use in this paper are Theorems 1 and 2 of [1]. If we translate the results of these theorems to our notation we obtain:

Lemma 4.1. *The function $y_k(t)$ defined by*

$$y_k(t)=\sum_{p=0}^{\infty}a_{pk}\Gamma(\mu_k+1-p)(t-\lambda_k)^{p-\mu_k-1}, \qquad |t-\lambda_k|<\min_{j\neq k}|\lambda_j-\lambda_k|, \tag{4.16}$$

is analytic in $\mathcal{P}_\eta$, satisfies

$$y_k(t)=\frac{K_{jk}}{1-e^{-2\pi i\mu_k}}y_j(t)+reg(t-\lambda_j), \qquad j\neq k, \tag{4.17}$$

where the K_{jk} are constants, and can be continued analytically along every path that does not intersect any of the points $\lambda_1,\cdots,\lambda_n$. Furthermore, if S is any sector in the t-plane of the form $S=\{|t|>R,\ \alpha<\mathrm{ph}t<\beta\}$ with $0<\beta-\alpha<2\pi$ and $R>\max|\lambda_j|$, then

$$\lim_{t\to\infty}e^{-(a+\varepsilon)|t|}y_k(t)=0, \qquad t\in S, \tag{4.18}$$

for $\varepsilon>0$ arbitrary.

In (4.17) $reg(t-\lambda_j)$ denotes a function that is regular (or analytic) in a neighbourhood of $t=\lambda_j$.

Lemma 4.2. *Let $\eta\in\mathbb{R}$ be admissible. If we define*

$$w_k(z,\eta)=\frac{1}{2\pi i}\int_{\gamma_k(\eta)}e^{zt}y_k(t)\,dt, \tag{4.19}$$

where $\gamma_k(\eta)$ is the contour in $\mathcal{P}_\eta$ from ∞ along the left-hand side of the cut $\mathrm{ph}(t-\lambda_k) = \eta$, *around λ_k in the positive sense, and back to ∞ along the right-hand side of the cut, then $w_k(z,\eta)$ is a solution of* (4.1), $w_k(z,\tilde{\eta}) = w_k(z,\eta)$ *for all $\tilde{\eta} \in \mathcal{I}_\eta$, and*

$$w_k(z,\eta) \sim e^{\lambda_k z} z^{\mu_k} \sum_{s=0}^{\infty} a_{sk} z^{-s}, \tag{4.20}$$

as $z \to \infty$ in $\mathcal{S}(\eta)$.

For each admissible η we have n solutions $w_1(z,\eta), \cdots, w_n(z,\eta)$. Since (4.1) is a linear ordinary differential equation of order n, for each admissible $\tilde{\eta}$ and $k \in \{1, \cdots, n\}$ there are connection coefficients $C_{jk}(\tilde{\eta},\eta)$ such that

$$w_k(z,\tilde{\eta}) = C_{1k}(\tilde{\eta},\eta) w_1(z,\eta) + \cdots + C_{nk}(\tilde{\eta},\eta) w_n(z,\eta). \tag{4.21}$$

If $\tilde{\eta} \in \mathcal{I}_\eta$, then $C_{jk}(\tilde{\eta},\eta) = \delta_{jk}$. Hence, the connection coefficients can change only when we cross a non admissible direction. The corresponding directions in the z-plane are generally known as Stokes lines. To compute all the connection coefficients it suffices to compute the connection coefficients of two neighbouring intervals $\mathcal{I}_\eta$.

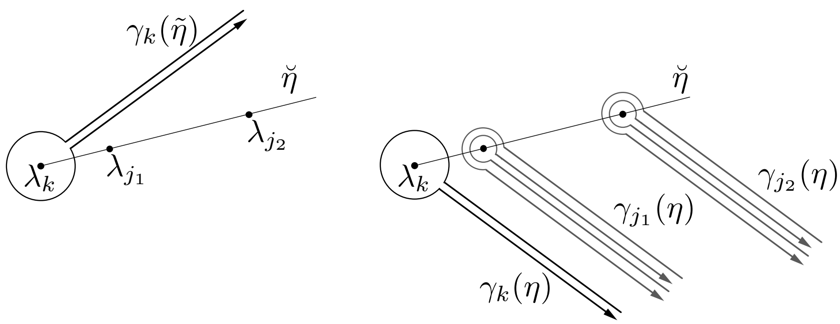


Fig. 7. $\gamma_k(\tilde{\eta})$ before (left) and after (right) the rotation.

Now assume that we are in the case of Figure 7. Then with (4.17) we obtain

$$\begin{aligned} w_k(z,\tilde{\eta}) &= w_k(z,\eta) + \sum_{l=1}^{p} \frac{1-e^{-2\pi i \mu_k}}{2\pi i} \int_{\gamma_{j_l}(\eta)} e^{zt} y_k(t)\,dt \\ &= w_k(z,\eta) + \sum_{l=1}^{p} \frac{K_{j_l k}}{2\pi i} \int_{\gamma_{j_l}(\eta)} e^{zt} y_{j_l}(t)\,dt \\ &= w_k(z,\eta) + \sum_{l=1}^{p} K_{j_l k} w_{j_l}(z,\eta). \end{aligned} \tag{4.22}$$

The constants K_{jk} are called the *Stokes multipliers*, and they play an important role in the definitions of the following numbers. Let

$$\begin{aligned}\alpha_k^{(m)} = \min\{|\lambda_k - \lambda_{j_0}| + |\lambda_{j_0} - \lambda_{j_1}| + \cdots + |\lambda_{j_{m-1}} - \lambda_{j_m}| \\ \mid j_0 \neq k, K_{j_0 k} \neq 0, j_l \neq j_{l-1}, K_{j_l j_{l-1}} \neq 0\}.\end{aligned} \tag{4.23}$$

If $G = (V, E)$ is a directed graph with edges $E = \{(\lambda_p, \lambda_q) \mid 1 \leq p, q \leq n, p \neq q,\ K_{qp} \neq 0\}$ and vertices $V = \{\lambda_1, \cdots, \lambda_n\}$, then $\alpha_k^{(m)}$ is the length of the shortest directed path of $m+1$ steps starting at λ_k. The main step needed to reach subsequent levels in the hyperasymptotic expansion for solutions of (4.1) is the following version of Taylor's theorem.

Lemma 4.3. *Let C be a closed contour encircling t and λ_k such that λ_j, $j \neq k$, is in the exterior of C. Then*

$$\begin{aligned}y_k(t) = \sum_{p=0}^{N-1} a_{pk} \Gamma(\mu_k + 1 - p)(t - \lambda_k)^{p-\mu_k-1} \\ + \frac{(t-\lambda_k)^{N-\mu_k-1}}{2\pi i} \int_C \frac{y_k(\tau)(\tau - \lambda_k)^{\mu_k+1-N}}{\tau - t}\, d\tau.\end{aligned} \tag{4.24}$$

We finish this section with the definition of *hyperterminants*. In the definition we shall use the notation

$$\int_\lambda^{[\eta]} = \int_\lambda^{\infty e^{i\eta}}, \qquad \eta \in \mathbb{R}. \tag{4.25}$$

Let l be a nonnegative integer, $\mathrm{Re} M_j > 1$, $\sigma_j \in \mathbb{C}$, $\sigma_j \neq 0$, $j = 0, \cdots, l$. Then

$$\begin{aligned}F^{(0)}(z) &= 1, \\ F^{(1)}\left(z; \begin{matrix} M_0 \\ \sigma_0 \end{matrix}\right) &= \int_0^{[\pi-\theta_0]} \frac{e^{\sigma_0 t_0} t_0^{M_0-1}}{z - t_0}\, dt_0, \\ F^{(l+1)}\left(z; \begin{matrix} M_0 \cdots, M_l \\ \sigma_0, \cdots, \sigma_l \end{matrix}\right) &= \\ \int_0^{[\pi-\theta_0]} \cdots \int_0^{[\pi-\theta_l]} & \frac{e^{\sigma_0 t_0 + \cdots + \sigma_l t_l} t_0^{M_0-1} \cdots t_l^{M_l-1}}{(z-t_0)(t_0-t_1)\cdots(t_{l-1}-t_l)}\, dt_l \cdots dt_0,\end{aligned} \tag{4.26}$$

where $\theta_j = \mathrm{ph}\sigma_j$, $j = 0, 1, \cdots, l$. The multiple integrals converge when $-\pi - \theta_0 < \mathrm{ph} z < \pi - \theta_0$.

4.3 The re-expansions

To obtain an integral representation for the remainder we substitute (4.24) into (4.19) and obtain

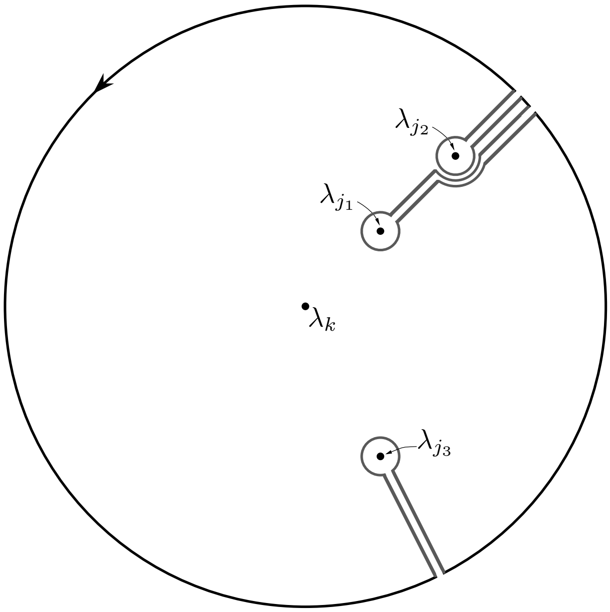


Fig. 8. Contour C in Lemma 4.3.

$$w_k(z,\eta) = e^{\lambda_k z} z^{\mu_k} \sum_{s=0}^{N_0-1} a_{sk} z^{-s} + R_k^{(0)}(z,\eta;N_0), \tag{4.27}$$

with

$$\begin{aligned} R_k^{(0)}(z,\eta;N_0) &= \frac{1}{(2\pi i)^2} \int_{\gamma_k(\eta)} \int_C e^{zt} \left(\frac{t-\lambda_k}{\tau-\lambda_k}\right)^{N_0-\mu_k-1} \frac{y_k(\tau)}{\tau - t}\, d\tau\, dt \\ &= \frac{e^{-2\pi i \mu_k}-1}{(2\pi i)^2} \int_{\lambda_k}^{[\eta]} \int_C e^{zt} \left(\frac{t-\lambda_k}{\tau-\lambda_k}\right)^{N_0-\mu_k-1} \frac{y_k(\tau)}{\tau - t}\, d\tau\, dt, \end{aligned} \tag{4.28}$$

where we collapsed $\gamma_k(\eta)$. Now take for C the contour given in Figure 8. In [8] it is shown that the contribution of the large circle in Figure 8 is negligible, and that when we replace the finite loops encircling λ_{j_p} by $\gamma_{j_p}(\theta_{kj_p})$ (which have the opposite orientation, note the extra minus sign in the next result) the error introduced is again negligible. In this document we will ignore negligible terms and write

$$\begin{aligned} &R_k^{(0)}(z,\eta;N_0)\\ &=\sum_{j\neq k}\frac{1-e^{-2\pi i\mu_k}}{(2\pi i)^2}\int_{\lambda_k}^{[\eta]}\int_{\gamma_j(\theta_{kj})}e^{zt}\left(\frac{t-\lambda_k}{\tau-\lambda_k}\right)^{N_0-\mu_k-1}\frac{y_k(\tau)}{\tau-t}\,d\tau\,dt\\ &=\sum_{j\neq k}\frac{K_{jk}}{(2\pi i)^2}\int_{\lambda_k}^{[\eta]}\int_{\gamma_j(\theta_{kj})}e^{zt}\left(\frac{t-\lambda_k}{\tau-\lambda_k}\right)^{N_0-\mu_k-1}\frac{y_j(\tau)}{\tau-t}\,d\tau\,dt, \end{aligned} \tag{4.29}$$

where we have used (4.17).

These are the main steps to get the re-expansions. Hence, for the next level we substitute (4.24) with $k=j$ in the final integral of (4.29). We obtain a finite sum of double integrals that can be expressed in terms of the first hyperterminant (see exercises). The loop $\gamma_j(\theta_{kj})$ is again collapsed, and we manipulate the new contour C in the same way as described in the previous paragraph. The result is

$$\begin{aligned} &R_k^{(0)}(z,\eta;N_0)=\\ &e^{\lambda_k z}z^{\mu_k+1-N_0}\sum_{j\neq k}\frac{K_{jk}}{2\pi i}\sum_{s=0}^{N_j^{(1)}-1}a_{sj}F^{(1)}\left(z;\begin{matrix}N_0-s+\mu_{jk}\\ \lambda_{jk}\end{matrix}\right)+R_k^{(1)}(z,\eta), \end{aligned} \tag{4.30}$$

where

$$\begin{aligned} &R_k^{(1)}(z,\eta)=\sum_{j\neq k}\sum_{l\neq j}\frac{K_{jk}K_{lj}}{(2\pi i)^3}\\ &\times\int_{\lambda_k}^{[\eta]}\int_{\lambda_j}^{[\theta_{kj}]}\int_{\gamma_l(\theta_{jl})}\frac{e^{zt_0}\left(\frac{t_0-\lambda_k}{t_1-\lambda_k}\right)^{N_0-\mu_k-1}\left(\frac{t_1-\lambda_j}{t_2-\lambda_j}\right)^{N_j^{(1)}-\mu_j-1}y_l(t_2)}{(t_1-t_0)(t_2-t_1)}\,dt_2\,dt_1\,dt_0. \end{aligned} \tag{4.31}$$

Note that the final integral line in (4.29) and (4.31) are very similar. Hence we can continue this process and obtain:

$$
\begin{aligned}
w_k(z,\eta) &= e^{\lambda_k z} z^{\mu_k} \Bigg(\sum_{s=0}^{N_k^{(0)}-1} a_{sk} z^{-s} + z^{1-N_k^{(0)}} \Bigg[\sum_{k_1 \neq k} \frac{K_{k_1 k}}{2\pi i} \\
&\times \Bigg\{ \sum_{s=0}^{N_{k_1}^{(1)}-1} a_{sk_1} F^{(1)} \left(z; \begin{matrix} N_k^{(0)} - s + \mu_{k_1 k} \\ \lambda_{k_1 k} \end{matrix} \right) + \\
&\sum_{k_2 \neq k_1} \frac{K_{k_2 k_1}}{2\pi i} \Bigg\{ \sum_{s=0}^{N_{k_2}^{(2)}-1} a_{sk_2} F^{(2)} \left(z; \begin{matrix} N_k^{(0)} - N_{k_1}^{(1)} + \mu_{k_1 k} + 1, & N_{k_1}^{(1)} - s + \mu_{k_2 k_1} \\ \lambda_{k_1 k}, & \lambda_{k_2 k_1} \end{matrix} \right) \\
&\quad \ddots \\
&\quad + \sum_{k_l \neq k_{l-1}} \frac{K_{k_l k_{l-1}}}{2\pi i} \Bigg\{ \sum_{s=0}^{N_{k_l}^{(l)}-1} a_{sk_l} F^{(l)} \Bigg(z; \begin{matrix} N_k^{(0)} - N_{k_1}^{(1)} + \mu_{k_1 k} + 1, \cdots, \\ \lambda_{k_1 k}, \qquad \cdots, \end{matrix} \\
&\qquad \begin{matrix} N_{k_{l-2}}^{(l-2)} - N_{k_{l-1}}^{(l-1)} + \mu_{k_{l-1} k_{l-2}} + 1, & N_{k_{l-1}}^{(l-1)} - s + \mu_{k_l k_{l-1}} \\ \lambda_{k_{l-1} k_{l-2}}, & \lambda_{k_l k_{l-1}} \end{matrix} \Bigg) \\
&\qquad \Bigg\} \cdots \Bigg\} \Bigg\} \Bigg] \Bigg) + R_k^{(l)}(z,\eta).
\end{aligned}
\tag{4.32}
$$

4.4 The optimal numbers of terms

The main result that is missing in the previous section is estimates for the remainders. Like in (4.29) and (4.31) it is possible to obtain an integral representation for $R_k^{(l)}(z,\eta)$ in (4.32). When we ignore the denominator in this integral representation, the integrals decouple in a product of gamma functions and beta integrals. The final result is:

Theorem 4.1. *Let l be an arbitrary nonnegative integer and $N_k^{(0)}$, $N_{k_1}^{(1)}, \cdots,$ $N_{k_l}^{(l)}$ be integers such that*

$$
N_k^{(0)} = \beta_k^{(0)}|z| + \gamma_k^{(0)}, \qquad N_{k_j}^{(j)} = \beta_{k_j}^{(j)}|z| + \gamma_{k_j}^{(j)}, \qquad j = 1, 2, \cdots, l, \tag{4.33}
$$

in which the β's are constants that satisfy

$$
0 < \beta_{k_l}^{(l)} < \beta_{k_{l-1}}^{(l-1)} < \cdots < \beta_{k_1}^{(1)} < \beta_k^{(0)} \tag{4.34}
$$

and the γ's are bounded as $|z| \to \infty$. Then as $z \to \infty$ in $\overline{\mathcal{S}}(\eta)$,

$$
\begin{aligned}
R_k^{(l)}(z,\eta) &= e^{\lambda_k z}\sum_{k_1\neq k}\cdots\sum_{k_{l+1}\neq k_l} K_{k_1k}\cdots K_{k_{l+1}k_l}|z|^{Re\mu_{k_{l+1}}+\frac{l+1}{2}}\\
&\times\left(\frac{\beta_k^{(0)}-\beta_{k_1}^{(1)}}{|\lambda_{k_1k}|e}\right)^{(\beta_k^{(0)}-\beta_{k_1}^{(1)})|z|}\cdots\left(\frac{\beta_{k_{l-1}}^{(l-1)}-\beta_{k_l}^{(l)}}{|\lambda_{k_lk_{l-1}}|e}\right)^{(\beta_{k_{l-1}}^{(l-1)}-\beta_{k_l}^{(l)})|z|}\\
&\times\left(\frac{\beta_{k_l}^{(l)}}{|\lambda_{k_{l+1}k_l}|e}\right)^{\beta_{k_l}^{(l)}|z|}\mathcal{O}(1).
\end{aligned}
\tag{4.35}
$$

Now what are the optimal choices for $N_k^{(0)},N_{k_1}^{(1)},\cdots,N_{k_l}^{(l)}$? Before we discus the general result let us first look at the simple example in which we take $n=2$. Hence, there are only two λ's. We also take $|\lambda_{12}|=|\lambda_1-\lambda_2|=1$ and $K_{12}\neq 0\neq K_{21}$. It seems to make sense to optimise $N_k^{(0)}$ in the level 0 expansion like we did in §2. Hence, we have to choose $\beta_k^{(0)}$ such that $\left(\beta_k^{(0)}/e\right)^{\beta_k^{(0)}|z|}$ is minimal. The optimal choice is $\beta_k^{(0)}=1$, that is $N_k^{(0)}\sim|z|$ as $|z|\to\infty$. Now suppose that we have optimised $N_k^{(0)},N_{k_1}^{(1)},\cdots,N_{k_{l-1}}^{(l-1)}$ in the previous levels and that we find $\beta_{k_l}^{(l)}$ by minimising

$$
\left(\frac{\beta_{k_{l-1}}^{(l-1)}-\beta_{k_l}^{(l)}}{e}\right)^{(\beta_{k_{l-1}}^{(l-1)}-\beta_{k_l}^{(l)})|z|}\left(\frac{\beta_{k_l}^{(l)}}{e}\right)^{\beta_{k_l}^{(l)}|z|}
$$

as a function of $\beta_{k_l}^{(l)}$. The optimal choice is $\beta_{k_l}^{(l)}=\frac{1}{2}\beta_{k_{l-1}}^{(l-1)}$. In this way we obtain $\beta_{k_l}^{(l)}=2^{-l}$. With these choices (4.35) simplifies to

$$
R_k^{(l)}(z,\eta)=e^{\lambda_k z-|z|}2^{-(2-2^{1-l})|z|}|z|^{\mathrm{Re}\mu_{k_{l+1}}+\frac{l+1}{2}}\mathcal{O}(1). \tag{4.36}
$$

Hence, the exponential improvement is $e^{-|z|}2^{-(2-2^{1-l})|z|}$ and when we let $l\to\infty$ then $e^{-|z|}2^{-2|z|}=e^{-(2.386\cdots)|z|}$ seems to be the optimal exponential improvement that we can obtain.

However, when we optimise

$$
\left(\frac{\beta_k^{(0)}-\beta_{k_1}^{(1)}}{e}\right)^{(\beta_k^{(0)}-\beta_{k_1}^{(1)})|z|}\cdots\left(\frac{\beta_{k_{l-1}}^{(l-1)}-\beta_{k_l}^{(l)}}{e}\right)^{(\beta_{k_{l-1}}^{(l-1)}-\beta_{k_l}^{(l)})|z|}\left(\frac{\beta_{k_l}^{(l)}}{e}\right)^{\beta_{k_l}^{(l)}|z|}
$$

as a function of $\beta_{k_l}^{(l)},\beta_{k_{l-1}}^{(l-1)},\cdots,\beta_k^{(0)}$ then we find that we have to take

$$
\beta_k^{(0)}-\beta_{k_1}^{(1)}=\cdots=\beta_{k_{l-1}}^{(l-1)}-\beta_{k_l}^{(l)}=\beta_{k_l}^{(l)}=1.
$$

Hence, the optimal choice is $\beta_{k_m}^{(m)}=l-m+1$. With these choices (4.35) simplifies to

$$R_k^{(l)}(z,\eta) = e^{\lambda_k z-(l+1)|z|}|z|^{\mathrm{Re}\mu_{k_{l+1}}+\frac{l+1}{2}}\mathcal{O}(1). \tag{4.37}$$

Hence, we can do much better than (4.36). The general result is:

Theorem 4.2. *Let*

$$\begin{aligned} \beta_k^{(0)} = \alpha_k^{(l)}, \quad \beta_{k_1}^{(1)} &= \max\left(0, \beta_k^{(0)} - |\lambda_{k_1 k}|\right), \\ \cdots, \ \beta_{k_l}^{(l)} &= \max\left(0, \beta_{k_{l-1}}^{(l-1)} - |\lambda_{k_l k_{l-1}}|\right), \end{aligned} \tag{4.38}$$

where $\alpha_k^{(l)}$ *is defined by* (4.23)*. Then, as* $z\to\infty$ *in* $\overline{\mathcal{S}}(\eta)$*,*

$$R_k^{(l)}(z,\eta) = e^{\lambda_k z-\alpha_k^{(l)}|z|}|z|^{\tilde{\mu}+\frac{l+1}{2}}\mathcal{O}(1) \tag{4.39}$$

where $\tilde{\mu}$ *is defined in* (4.8)*.*

4.5 The computation of the Stokes multipliers

In the next section we will discus the computation of the hyperterminants. For hyperasymptotic expansion (4.32) to be useful we have to be able to compute the Stokes multipliers K_{jk}. The Stokes multipliers are defined in (4.22). It seems not to be possible to use (4.22) to compute these constants. Hence, we need an efficient method to compute these Stokes multipliers.

First we obtain another hyperasymptotic result. Note that the coefficients in the asymptotic expansions (4.3) can also be written as

$$a_{N_k^{(0)}k} = e^{-\lambda_k z} z^{N_k^{(0)}-\mu_k}\left(R^{(0)}(z,\eta;N_k^{(0)}) - R^{(0)}(z,\eta;N_k^{(0)}+1)\right). \tag{4.40}$$

This seems to be useless were it not that we have hyperasymptotic expansions for the remainders $R^{(0)}(z,\eta;N_k^{(0)}+m)$. We substitute these hyperasymptotic expansions and note that when we simplify the result z disappears:

$$
\begin{aligned}
&a_{N_k^{(0)}k} = \\
&-\sum_{k_1\neq k}\frac{K_{k_1k}}{2\pi i}\left\{\sum_{s=0}^{N_{k_1}^{(1)}-1} a_{sk_1}F^{(1)}\left(0;\,\begin{matrix}(N_k^{(0)}+1)-s+\mu_{k_1k}\\ \lambda_{k_1k}\end{matrix}\right)+\sum_{k_2\neq k_1}\frac{K_{k_2k_1}}{2\pi i}\right.\\
&\times\left\{\sum_{s=0}^{N_{k_2}^{(2)}-1} a_{sk_2}F^{(2)}\left(0;\,\begin{matrix}(N_k^{(0)}+1)-N_{k_1}^{(1)}+\mu_{k_1k}+1, & N_{k_1}^{(1)}-s+\mu_{k_2k_1}\\ \lambda_{k_1k}, & \lambda_{k_2k_1}\end{matrix}\right)\right.\\
&\quad\ddots\\
&+\sum_{k_l\neq k_{l-1}}\frac{K_{k_lk_{l-1}}}{2\pi i}\left\{\sum_{s=0}^{N_{k_l}^{(l)}-1} a_{sk_l}F^{(l)}\left(0;\,\begin{matrix}(N_k^{(0)}+1)-N_{k_1}^{(1)}+\mu_{k_1k}+1, & \cdots,\\ \lambda_{k_1k}, & \cdots,\end{matrix}\right.\right.\\
&\qquad\left.\begin{matrix}N_{k_{l-2}}^{(l-2)}-N_{k_{l-1}}^{(l-1)}+\mu_{k_{l-1}k_{l-2}}+1, & N_{k_{l-1}}^{(l-1)}-s+\mu_{k_lk_{l-1}}\\ \lambda_{k_{l-1}k_{l-2}}, & \lambda_{k_lk_{l-1}}\end{matrix}\right)\\
&\qquad\qquad\left.\left.\left.\right\}\cdots\right\}\right\}+r_k^{(l)}(N_k^{(0)}).
\end{aligned}
\tag{4.41}
$$

This may be regarded as a hyperasymptotic expansion for the late coefficients as $N_k^{(0)}\to\infty$. In order to use this result for the computation of the Stokes multipliers, we suppose that the numbers of coefficients are given by (4.33). Thus, instead of $N_k^{(0)}$, we take $|z|$ as the large parameter.

It follows from the previous section that if (4.38) holds, then

$$
r_k^{(l)}(N_k^{(0)}) = e^{-\alpha_k^{(l)}|z|}|z|^{N_k^{(0)}-\mu_k+\tilde{\mu}+\frac{l+1}{2}}\mathcal{O}(1), \tag{4.42}
$$

as $|z|\to\infty$.

The main difference between (4.32) and (4.41) is that in (4.41) the only unknowns are the Stokes multipliers. On replacing in (4.41) $N_k^{(0)}$ by $N_k^{(0)}-1, N_k^{(0)}-2,\cdots,N_k^{(0)}-n_1$, in turn, and ignoring the error terms, we arrive at a system of n_1 equations. In this way we can compute the $K_{k_jk_m}$ to the required precision. More details are given in [8] and [9].

4.6 The computation of the hyperterminants

In [6] and [7] (see also the exercises) it is shown that we can compute the hyperterminants via

$$
\begin{aligned}
&F^{(l+1)}\left(z;\,\begin{matrix}M_0,\cdots,M_l\\ \sigma_0,\cdots,\sigma_l\end{matrix}\right) = \\
&\qquad\sum_{k=0}^{\infty}A^{(l+1)}\left(k;\,\begin{matrix}M_0,\cdots,M_l\\ \sigma_0,\cdots,\sigma_l\end{matrix}\right)U(k+1,2-M_0,z\sigma_0),
\end{aligned}
\tag{4.43}
$$

where the coefficients $A^{(m)}(k;\cdots)$ can be computed via

$$A^{(1)}\left(k;\begin{matrix}M_0\\ \sigma_0\end{matrix}\right)=\delta_{k0}\,\sigma_0^{1-M_0}\,e^{M_0\pi i}\,\Gamma(M_0), \tag{4.44}$$

and

$$\begin{aligned}A^{(l+1)}\left(k;\begin{matrix}M_0,\ \cdots,\ M_l\\ \sigma_0,\ \cdots,\ \sigma_l\end{matrix}\right)&=e^{\pi i(M_0+k)}\,\sigma_0^{1-M_0-k}\,\sigma_1^k\,\Gamma(M_0+M_1-1)\\ \times\,\Gamma(M_0+k)\sum_{s=0}^{\infty}&\left[\frac{(k+s)!}{s!\Gamma(M_0+M_1+k+s)}\right.\\ &\times\,{}_2F_1(M_0+k,k+s+1;M_0+M_1+k+s;1+\frac{\sigma_1}{\sigma_0})\\ &\left.\times A^{(l)}\left(s;\begin{matrix}M_1,\ \cdots,\ M_l\\ \sigma_1,\ \cdots,\ \sigma_l\end{matrix}\right)\right].\end{aligned} \tag{4.45}$$

In (4.43) $U(a,c,z)$ is the confluent hypergeometric U-function.

4.7 Example

We use the example

$$\begin{aligned}w^{(4)}(z)-3w^{(3)}(z)+\left(\frac{9}{4}+\frac{1}{2}z^{-2}\right)w^{(2)}(z)-\left(3+\frac{3}{4}z^{-2}\right)w'(z)&\\ +\left(\frac{5}{4}+\frac{9}{16}z^{-2}\right)w(z)=0.&\end{aligned} \tag{4.46}$$

In this case we have $n=4$ and

$$\lambda_1=\frac{1}{2},\quad \lambda_2=\frac{5}{2},\quad \lambda_3=i,\quad \lambda_4=-i,\quad \mu_j=0,\quad j=1,\cdots,4. \tag{4.47}$$

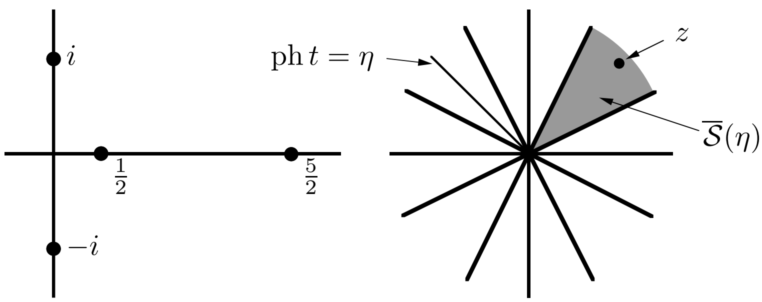


Fig. 9. The admissible directions.

We take $\eta = \frac{3}{4}\pi$ and assume that we wish to compute $w_3(z,\eta)$ at $z = 10e^{\frac{1}{4}\pi i}$. Since,

$$\lambda_{13} = \frac{1}{2} - i, \qquad \lambda_{23} = \frac{5}{2} - i, \qquad \lambda_{43} = -2i, \tag{4.48}$$

we obtain from (4.23) $\alpha_3^{(0)} = |\lambda_{13}| = \frac{1}{2}\sqrt{5} = 1.118\cdots$. Hence, the optimal number of terms in the original Poincaré expansion is $N_3^{(0)} = 11$.

In the level one expansion we have $\beta_3^{(0)} = \alpha_3^{(1)} = |\lambda_{13}| + |\lambda_{31}| = |\lambda_{13}| + |\lambda_{41}| = \sqrt{5} = 2.236\cdots$. Hence,

$$\beta_3^{(0)} = \sqrt{5} = 2.236\cdots, \begin{cases} \beta_1^{(1)} = \max\left(0, \beta_3^{(0)} - |\lambda_{13}|\right) = \frac{1}{2}\sqrt{5} = 1.118\cdots, \\ \beta_2^{(1)} = \max\left(0, \beta_3^{(0)} - |\lambda_{23}|\right) = 0, \\ \beta_4^{(1)} = \max\left(0, \beta_3^{(0)} - |\lambda_{43}|\right) = \sqrt{5} - 2 = 0.236\cdots. \end{cases} \tag{4.49}$$

Thus the optimal numbers of terms at the level one hyperasymptotic expansion are $N_3^{(0)} = 22$, $N_1^{(1)} = 11$, $N_2^{(1)} = 0$ and $N_4^{(1)} = 2$.

Since $\beta_2^{(1)} = 0$ we don't have to compute K_{23} at this level. To compute the Stokes multipliers K_{13} and K_{43} to the required precision, we first compute

$$a_{N_3^{(0)}-1,3} = (-2.62461157451487375 38\cdots + i7.73814877018878589 60\cdots)10^{16},$$
$$a_{N_3^{(0)}-2,3} = (-3.21453466302543406 02\cdots - i3.24638461670627105 32\cdots)10^{15}.$$

On replacing $N_k^{(0)}$ by $N_3^{(0)} - 1$, and $N_3^{(0)} - 2$ in turn in the level one version of (4.41), and ignoring the error terms, we arrive at a system of 2 linear equations. On solving these equations we find that

$$\begin{aligned} K_{13} &= -1.3167355555044700 9754 + 1.7502706941028591333i, \\ K_{43} &= 0.3443668989089293253 - 2.1203898698716294832i. \end{aligned} \tag{4.50}$$

Further on we shall see that in these approximations the value of K_{13} is correct to 8 decimal places and that of K_{43} is correct to 2 decimal places. The required precision for K_{13} is approximately $\left(\beta_3^{(0)}/|\lambda_{13}|\right)^{-\beta_3^{(0)}|z|} = 1.9 \times 10^{-7}$, and the required precision for K_{43} is approximately $\left(\beta_3^{(0)}/|\lambda_{43}|\right)^{-\beta_3^{(0)}|z|} = 8.3 \times 10^{-2}$.

In the level two expansion we have $\beta_3^{(0)} = \alpha_3^{(2)} = \frac{3}{2}\sqrt{5}$. Hence,

$$\beta_3^{(0)} = 3.354\cdots, \quad \begin{cases} \beta_1^{(1)} = 2.236\cdots, & \begin{cases} \beta_2^{(2)} = 0.236\cdots, \\ \beta_3^{(2)} = 1.118\cdots, \\ \beta_4^{(2)} = 1.118\cdots, \end{cases} \\ \beta_2^{(1)} = 0.661\cdots, & \quad \beta_j^{(2)} = 0, \quad j = 1,3,4, \\ \beta_4^{(1)} = 1.354\cdots, & \begin{cases} \beta_1^{(2)} = 0.236\cdots, \\ \beta_j^{(2)} = 0, \quad j = 2,3. \end{cases} \end{cases} \tag{4.51}$$

Thus we take

$$N_3^{(0)} = 33, \quad \begin{cases} N_1^{(1)} = 22, & \begin{cases} N_2^{(2)} = 2, \\ N_3^{(2)} = 11, \\ N_4^{(2)} = 11, \end{cases} \\ N_2^{(1)} = 6, & \quad N_j^{(2)} = 0, \quad j = 1, 3, 4, \\ N_4^{(1)} = 13, & \begin{cases} N_1^{(2)} = 2, \\ N_j^{(2)} = 0, \quad j = 2, 3. \end{cases} \end{cases} \tag{4.52}$$

Hence the level two hyperasymptotic expansion of $w_3(z,\eta)$, with $z = 10e^{\frac{1}{4}\pi i}$, is of the form

$$\begin{aligned} e^{-iz} w_3(z,\eta) = & \sum_{s=0}^{32} \frac{a_{s3}}{z^s} + z^{-32} \frac{K_{13}}{2\pi i} \sum_{s=0}^{21} a_{s1} F^{(1)} \left(z; \begin{matrix} 33 - s \\ \frac{1}{2} - i \end{matrix} \right) \\ & + z^{-32} \frac{K_{23}}{2\pi i} \sum_{s=0}^{5} a_{s2} F^{(1)} \left(z; \begin{matrix} 33 - s \\ \frac{5}{2} - i \end{matrix} \right) + z^{-32} \frac{K_{43}}{2\pi i} \sum_{s=0}^{12} a_{s4} F^{(1)} \left(z; \begin{matrix} 33 - s \\ -2i \end{matrix} \right) \\ & + z^{-32} \frac{K_{13}}{2\pi i} \left[\frac{K_{21}}{2\pi i} \sum_{s=0}^{1} a_{s2} F^{(2)} \left(z; \begin{matrix} 12, \ 22 - s \\ \frac{1}{2} - i, \ 2 \end{matrix} \right) + \right. \\ & \left. \frac{K_{31}}{2\pi i} \sum_{s=0}^{10} a_{s3} F^{(2)} \left(z; \begin{matrix} 12, \ 22 - s \\ \frac{1}{2} - i, \ i - \frac{1}{2} \end{matrix} \right) + \frac{K_{41}}{2\pi i} \sum_{s=0}^{10} a_{s4} F^{(2)} \left(z; \begin{matrix} 12, \ 22 - s \\ \frac{1}{2} - i, \ -i - \frac{1}{2} \end{matrix} \right) \right] \\ & + z^{-32} \frac{K_{43}}{2\pi i} \frac{K_{14}}{2\pi i} \sum_{s=0}^{1} a_{s1} F^{(2)} \left(z; \begin{matrix} 21, \ 13 - s \\ -2i, \ i + \frac{1}{2} \end{matrix} \right) + R_3^{(2)}(z,\eta). \end{aligned} \tag{4.53}$$

To compute the Stokes multipliers K_{21}, K_{31} and K_{41} we use the level one hyperasymptotic expansions of $a_{N_1^{(1)}-1,1}$, $a_{N_1^{(1)}-2,1}$ and $a_{N_1^{(1)}-3,1}$. We obtain

$$\begin{aligned} K_{21} &= 0.32220037911218913862i, \\ K_{31} &= -0.33518471810856053233 - 0.17394369472610190908i, \\ K_{41} &= 0.33518471810856053233 - 0.17394369472610190908i. \end{aligned} \tag{4.54}$$

The level one hyperasymptotic expansion of $a_{N_4^{(1)}-1,4}$ yields

$$K_{14} = 1.3175812208411643253 + 1.7492444366777255110i. \tag{4.55}$$

If we use (4.54) and (4.55) in the level two hyperasymptotic expansions of $a_{N_3^{(0)}-1,3}$, $a_{N_3^{(0)}-2,3}$ and $a_{N_3^{(0)}-3,3}$, then we obtain

$$\begin{aligned} K_{13} &= -1.3167355300409799821 + 1.7502707419178753228i, \\ K_{23} &= -0.9600011769004704206 - 0.3257656311181001083i, \\ K_{43} &= 0.3553405998176582756 - 2.1172377431478990710i. \end{aligned} \tag{4.56}$$

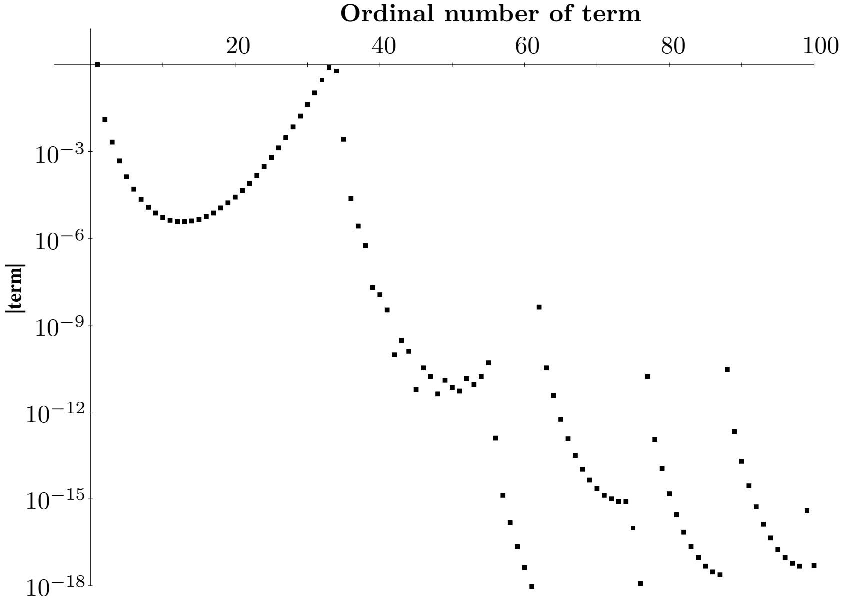


Fig. 10. The terms in (4.53), the level two hyperasymptotic expansion of $w_3(z,\eta)$. The first (U-shaped) group of terms are the terms in the original asymptotic expansion. The next three groups are the level one re-expansions, and the final four groups are the level two re-expansions.

To compute the level two hyperasymptotic expansion of $w_3(z,\eta)$, we use these values for the Stokes multipliers in (4.53) and the methods described in §4.6 to compute the hyperterminants. To compute the 'exact' value of $w_3(z,\eta)$ at $z = 10e^{\frac{1}{2}\pi i}$, we first use 47 terms of the asymptotic expansion of $w_3(z,\eta)$ to compute this function at the point $z+35i$, and then use numerical integration of the differential equation (4.46) from $z+35i$ to z. For more details on the stability of the numerical integration process see [10]. The numerical results are given in Table 1.

level	approximation	|relative error|
0	0.00059870442702324531293 $+0.00059285912251762980202i$	1.9×10^{-6}
1	0.00059870426695596695104 $+0.00059286076905849279426i$	1.7×10^{-13}
2	0.00059870426695611033390 $+0.00059286076905847214738i$	1.9×10^{-19}
exact	0.00059870426695611033376 +0.00059286076905847214746i	0

Table 1. Hyperasymptotic approximations to $w_3(z,\eta)$ for $z = 10e^{\frac{1}{4}\pi i}$.

5 Exercises

1. Let $f(z)$ be the function defined in (2.1).

(a) Take $z=10$ and compute the first 16 digits of $f(z)$.

(b) Take $N_0=10$ and compute $T_1=\displaystyle\sum_{n=0}^{N_0-1}\frac{(-)^n n!}{z^{n+1}}$.

(c) Take $N_1=13$ and compute $T_2=T_1+\displaystyle\sum_{n=0}^{N_1-1}2^{-n-1}p_n(z,N_0)$.

(d) Take $N_2=25$ and compute

$$T_3=T_2+\sum_{n=0}^{N_2-1}4^{-n-1}\int_0^\infty e^{-zt}(-t)^{N_0}\left(\frac{1-t}{2}\right)^{N_1}(3-t)^n\,dt.$$

2. The confluent hypergeometric U-function has integral representation

$$U(a,c,z)=\frac{1}{\Gamma(a)}\int_0^\infty e^{-zt}t^{a-1}(1+t)^{c-a-1}\,dt.$$

(a) Use the Taylor series expansion $(1-x)^{-\alpha}=\displaystyle\sum_{n=0}^{\infty}\frac{(\alpha)_n}{n!}x^n$ to construct the asymptotic series for $U(a,c,z)$ as $z\to\infty$.

(b) What is the optimal number of terms in this divergent asymptotic expansion?

(c) Construct the first re-expansion.

3. Recall that on a steepest descent path $z\,(h(t)-\lambda_n)\le 0$. To prove integral representation (3.7) we use the substitution

$$u=z\,(\lambda_n-h(t))\,, \tag{5.1}$$

in (3.3). Note that for a fixed $u>0$ (5.1) has two solutions. We call them t_- and t_+, where contour $C_n(\theta)$ passes first through t_- and then through t_+. Show that the result of this substitution is

$$T^{(n)}(z)=\frac{1}{\sqrt{z}}\int_0^\infty e^{-u}\left(\frac{g(t_-)}{h'(t_-)}-\frac{g(t_+)}{h'(t_+)}\right)du.$$

Now check that

$$\frac{g(t_-)}{h'(t_-)}-\frac{g(t_+)}{h'(t_+)}=\frac{1}{2\pi i\sqrt{u}}\int_{\Gamma_n(\theta)}\frac{g(t)\sqrt{z(\lambda-h(t))}}{\lambda_n-h(t)-u/z}\,dt.$$

4. In this exercise we will illustrate how representation (3.9) can be used. We take $g(t)=1$, $h(t)$ the function defined in (3.4), and the λ_j are given in (3.6). To compute the coefficients a_{s4} via (3.9) we note that $\lambda_4-h(t)=(t+i)^2\big(i-3+(1-3i)(t+i)+(\frac{3}{4}+i)(t+i)^2-\frac{1}{5}(t+i)^3\big)$. Hence

$$
\begin{aligned}
&(\lambda_4 - h(t))^{-s-1/2} = \\
&(t+i)^{-2s-1}\big(i-3+(1-3i)(t+i)+(\frac{3}{4}+i)(t+i)^2-\frac{1}{5}(t+i)^3\big)^{-s-1/2}.
\end{aligned}
\tag{5.2}
$$

Thus a_{s4} is equal to a gamma function times the coefficient of $(t+i)^{2s}$ in the Taylor series expansion of the final factor of the right-hand side of (5.2). Use this result to compute $a_{0,4}$, $a_{1,4}$, $\cdots$, $a_{20,4}$.

Similarly, use

$$
(\lambda_1 - h(t))^{-s-1/2} = (t-1)^{-2s-1}\big(1-\frac{1}{4}(t-1)^2-\frac{1}{5}(t-1)^3\big)^{-s-1/2}
$$

to compute $a_{0,1}, a_{1,1}, \cdots, a_{5,1}$.

5. Check the step from (3.11) to (3.12).

6. Use

$$
a_{Nn} = z^N\left(R_n^{(0)}(z,N+1) - R_n^{(0)}(z,N)\right)
$$

and (3.12) to obtain the integral representation

$$
a_{Nn} = \sum_{\text{adjacent } m} \frac{(-)^{\gamma_{nm}}}{2\pi i}\int_0^{\infty/\lambda_{nm}} e^{-\lambda_{nm}t}t^{N-1}T^{(m)}(t)\,dt.
$$

Now expand $T^{(m)}(t)$ in an asymptotic expansion for large t. The result is (3.16). You can check this result numerically with the coefficients that you computed in question 4.

7. In the derivation of (4.30) we have to show that

$$
\begin{aligned}
\frac{\Gamma(\mu_j+1-s)}{2\pi i}\int_{\lambda_k}^{[\eta]}\int_{\gamma_j(\theta_{kj})} e^{zt}\left(\frac{t-\lambda_k}{\tau-\lambda_k}\right)^{N-\mu_k-1}\frac{(\tau-\lambda_j)^{s-\mu_j-1}}{\tau-t}\,d\tau\,dt \\
= e^{\lambda_k z}z^{\mu_k+1-N}F^{(1)}\left(z;\begin{matrix}N-s+\mu_{jk}\\ \lambda_{jk}\end{matrix}\right).
\end{aligned}
$$

To obtain this result use the substitutions

$$
t_0 = z\frac{t-\lambda_k}{\tau-\lambda_k}, \qquad t_1 = t_0(\tau-\lambda_j), \quad \text{and} \quad \frac{d\tau\,dt}{\tau-t} = \frac{dt_1\,dt_0}{t_0(z-t_0)},
$$

and Hankel's loop integral representation for the reciprocal gamma function.

8. In the derivation of (4.37) we have to show that

$$
F(y_1,\cdots,y_n) = y_1^{y_1}\cdots y_n^{y_n}e^{-y_1-\cdots-y_n},
$$

has a minimum at $y_1 = y_2 = \cdots = y_n = 1$. Show that this is the case.

9. In this exercise we produce new integral representations for the hyperterminants. We will assume that $|\mathrm{ph}z+\theta_0| \le \frac{1}{2}\pi$ and that $\mathrm{Re}M_j > 1$ for all j.

(a) In the definition

$$F^{(1)}\left(z; \frac{M_0}{\sigma_0}\right) = \int_0^{[\pi-\theta_0]} \frac{e^{\sigma_0 t_0} t_0^{M_0-1}}{z-t_0}\, dt_0$$

use $\dfrac{1}{z-t_0} = \displaystyle\int_0^{[\theta_0]} e^{-s_0(z-t_0)}\, ds_0$ and show that

$$F^{(1)}\left(z; \frac{M_0}{\sigma_0}\right) = \Gamma(M_0) \int_0^{[\theta_0]} e^{-zs_0} \left(-s_0-\sigma_0\right)^{-M_0}\, ds_0.$$

(b) Now assume for the moment that $\theta_j - \theta_{j-1} \in [\frac{1}{2}\pi, \frac{3}{2}\pi] \mod 2\pi$. Use the previous substitution and similar substitutions for the other factors in the denominator to show that

$$\begin{aligned} &F^{(l+1)}\left(z; \frac{M_0, \cdots, M_l}{\sigma_0, \cdots, \sigma_l}\right) \\ &= \int_0^{[\pi-\theta_0]} \cdots \int_0^{[\pi-\theta_l]} \frac{e^{\sigma_0 t_0 + \cdots + \sigma_l t_l} t_0^{M_0-1} \cdots t_l^{M_l-1}}{(z-t_0)(t_0-t_1)\cdots(t_{l-1}-t_l)}\, dt_l \cdots dt_0 \\ &= \prod_{j=0}^{l} \Gamma(M_j) \int_0^{[\theta_0]} \cdots \int_0^{[\theta_l]} e^{-zs_0} \left(s_1 - s_0 - \sigma_0\right)^{-M_0} \cdots \\ &\qquad \times \left(s_l - s_{l-1} - \sigma_{l-1}\right)^{-M_{l-1}} \left(-s_l - \sigma_l\right)^{-M_l}\, ds_l \cdots ds_0. \end{aligned}$$

Analytic continuation removes the restriction on $\theta_j - \theta_{j-1}$.

(c) Expansion (4.43) is produced via

$$\begin{aligned} &\left(s_1 - s_0 - \sigma_0\right)^{-M_0} \\ &= \sigma_0^{M_0} \left(s_0+\sigma_0\right)^{-M_0} \left(s_1-\sigma_0\right)^{-M_0} \left(1 - \frac{s_0 s_1}{(s_0+\sigma_0)(s_1-\sigma_0)}\right)^{-M_0} \\ &= \sigma_0^{M_0} \sum_{k=0}^{\infty} \frac{\Gamma(M_0+k)}{\Gamma(M_0)k!} s_0^k (s_0+\sigma_0)^{-M_0-k} s_1^k (s_1-\sigma_0)^{-M_0-k}, \end{aligned}$$

and noting that

$$\frac{1}{k!}\int_0^{[\theta_0]} e^{-zs_0} s_0^k (s_0+\sigma_0)^{-M_0-k}\, ds_0 = \sigma_0^{1-M_0} U(k+1, 2-M_0, z\sigma_0).$$

Give an integral representation for the coefficients $A^{(l+1)}(..)$ in (4.43).

(d) Use a similar expansion for $\left(s_2 - s_1 - \sigma_1\right)^{-M_1}$ to obtain 'recurrence' relation (4.45).

References

1. W. Balser, W. B. Jurkat and D. A. Lutz, *On the reduction of connection problems for differential equations with an irregular singular point to ones with only regular singularities I*, SIAM J. Math. Anal. **12** (1981), 691–721.

2. M. V. Berry, *Uniform asymptotic smoothing of Stokes's discontinuities*, Proc. Roy. Soc. London, Ser. A**422** (1989), 7–21.
3. M. V. Berry and C. J. Howls, *Hyperasymptotics for integrals with saddles*, Proc. Roy. Soc. London, Ser. A**434** (1991), 657–675.
4. A. B. Olde Daalhuis, *Hyperasymptotic expansions of confluent hypergeometric functions*, IMA J. Appl. Math. **49** (1992), 203–216.
5. ———, *Hyperasymptotics and the Stokes' phenomenon*, Proc. Roy. Soc. Edinburgh Sect. A**123** (1993), 731–743.
6. ———, *Hyperterminants I*, J. Comp. Appl. Math. **76** (1996), 255–264.
7. ———, *Hyperterminants II*, J. Comput. Appl. Math. **89** (1998), 87–95.
8. ———, *Hyperasymptotic solutions of higher order differential equations with a singularity of rank one*, Proc. Roy. Soc. London, Ser. A**454** (1998), 1–29.
9. ———, *On the computation of Stokes multipliers via hyperasymptotics*, Sūrikaisekikenkyūsho Kōkyūroku **1088**, (1999), 68–78.
10. A. B. Olde Daalhuis and F. W. J. Olver, *On the asymptotic and numerical solution of linear ordinary differential equations*, SIAM Rev. **40**, (1998), 463–495.
11. F. W. J. Olver, *Asymptotics and Special Functions*, Academic Press, Inc., Boston, 1974. Reprinted by AK Peters, Wellesley, 1997.
12. H. Segur, S. Tanveer, and H. Levine, *Asymptotics beyond all orders*, NATO Advanced Science Institutes Series B: Physics, Vol. 284, Plenum Press, New York, 1991.
13. G. G. Stokes, *On the discontinuity of arbitrary constants which appear in divergent developments*, Trans. Camb. Phil. Soc. **10** (1857), 106–128, reprinted in "Mathematical and Physical Papers by the late Sir George Gabriel Stokes", CUP, 1905, vol IV, 77–109.
14. G. G. Stokes, *On the discontinuity of arbitrary constants that appear as multipliers of semi-convergent series*, Acta Math. **26** (1902), 393–397, reprinted in "Mathematical and Physical Papers by the late Sir George Gabriel Stokes", CUP, 1905, vol V, 283–287.
15. N. M. Temme, *Special Functions: An introduction to the Classical Functions of Mathematical Physics*, Wiley, New York, 1996.
16. R. Wong, *Asymptotic Approximations of Integrals*, Academic Press, New York, 1989.

Index

$f(z) = \mathcal{O}(g(z))$, 214
$f(z) = o(g(z))$, 214
$f(z) \sim \sum_{s=0}^{\infty} a_s z^{-s}$, 214
$f(z) \sim g(z)$, 214
$\star$-algebra, 30
$\star$-representation, 31
$3nj$-coefficient, 72
$6j$-coefficient, 59
$9j$-coefficient, 88

adjacent saddles, 221
adjoint representation, 35
admissible, 227
Airy functions, 168, 197
alternating sign matrix, 164
anti-symmetry, 27
Askey-Gasper identity, 11
Askey-Wilson integral, 157
Askey-Wilson scheme, 13
asymptotic expansion, 169, 182, 203, 214
asymptotic variable, 214

backward difference operator, 21
basic hypergeometric series, 19, 138
Bateman integral representation, 16
Bessel functions, 18, 197, 204
Bessel system, 107
Bieberbach conjecture, 11
Biedenharn-Elliott identity, 75, 81
big q-Legendre polynomials, 19, 22
binary coupling scheme, 71
binary coupling tree, 77
binary tree, 72
Borwein conjecture, 165
bracket, 27

Casimir operator, 28, 32, 33
Catalan numbers, 157
Cauchy transform, 170, 178, 180
Cauchy's formula, 169
Cauchy's integral formula, 17
character, 141, 162
characteristic equation, 226
Charlier polynomials, 158
Chebyshev polynomials, 151, 157, 195
Chu-Vandermonde formula, 11, 13, 37
classical q-orthogonal polynomials, 22
classical discrete orthogonal polynomials, 21
classical orthogonal polynomials, 20
Clausen's formula, 10
Clebsch-Gordan coefficients for the positive discrete series, 53
Clebsch-Gordan coefficients of $\mathfrak{su}(2)$, 44
Clebsch-Gordan coefficients, orthogonality relations, 48
`closedform`, 9
column strict tableau, 162, 164
commutator, 27
complete hyperasymptotic expansion, 226
complete matchings, 158
completely reducible $\mathfrak{g}$-module, 29
content numbers, 163
`contgosper`, 15
continued fraction, 159

continuous q-Hermite polynomials, 150, 158
continuous boundary values, 170
continuous Hahn polynomial, 81
continuous Hahn polynomial, convolution identity, 83
contraction semigroup, 121
convolution identity, 82–84
coupled basis, 48
crossing number, 158
crystallographic root system, 95

definite integration, 15
definite summation, 7
degree-lowering operator, 118
diameter, 80
dihedral group, 96
Dougall's identity, 9
dual Hahn polynomial, 51
dual Hahn polynomial, recurrence relation, 52
Dunkl kernel, 105
Dunkl kernel, asymptotics, 129
Dunkl kernel, integral representation, 120
Dunkl Laplacian, 100
Dunkl operator, 97
Dunkl transform, 109
Dunkl type heat operator, 122
Durfee square, 146

endpoint conditions, 180
enveloping algebra, 28
Euler pentagonal number theorem, 146
Euler's transformation formula, 36

`fasenmyer`, 6
Fasenmyer's method, 6
Feller-Markov semigroup, 124
Ferrers diagram, 138, 162
Fischer product, 100
forward difference operator, 8
FPS (formal power series), 3

$\mathfrak{g}$-module, 29
$\mathfrak{g}$-module, completely reducible, 29
$\mathfrak{g}$-module, irreducible, 29
$\mathfrak{g}$-module, simple, 29
$\mathfrak{g}$-modules, tensor product of, 29
$\mathfrak{g}$-submodule, 29
Gauss identity, 11
Gaussian kernel, 121
generalized translation, 122
generalized Bessel function, 107
generalized heat kernel, 123
generalized Hermite functions, 115
generalized Hermite polynomials, 115
generalized hypergeometric series, 37
generating functions, 17
generator of a strongly continuous semigroup, 121
`GFdiffeq`, 17
`GFrecursion`, 17
gfun, 4
$\mathfrak{gl}(n, \mathbb{C})$, 27
global hyperasymptotics, 221, 226
`gosper`, 8
Gosper's algorithm, 8, 12

Hahn polynomial, 50
Hankel matrix, 159
harmonic numbers, 8
Hermite functions, generalized, 115
Hermite polynomials, 115, 150, 158
Hermite polynomials, generalized, 115
highest weight vector, 32
holonomic differential equation, 4
holonomic function, 4
holonomic sequence, 4
`HolonomicDE`, 4
homomorphism, 28
hook numbers, 163
hook-content formula, 163
hsols, 14
hsum, 3, 5, 6, 8, 9, 13–15, 17
hyperasymptotic expansion, 219, 220, 223, 235, 236
hyperasymptotics, global, 221, 226
hyperasymptotics, local, 217
hyperexponential sums, 13
`hypergeom`, 3
hypergeometric series, 3, 35
hypergeometric series, generalized, 37
hypergeometric term, 3
`hyperterm`, 3
hyperterminant, 225, 230, 235, 236

implied constant, 214

indefinite integration, 15
indefinite summation, 8
`infclosedform`, 12
infhsum, 12
`intdiffeq`, 15
integer partition, 138
intermediate coupling, 57
internal node, 72
intertwining operator, 103
intertwining operator, positivity, 117
`intrecursion`, 15
involution, 149
involution principle, 150
irreducible $\star$-representation of $\mathfrak{su}(1,1)$, 33
irreducible $\mathfrak{g}$-module, 29
irreducible root system, 97
isotropic harmonic oscillator, 111, 112

Jack polynomials, non-symmetric, 116
Jacobi polynomials, 14, 17, 81, 182, 208
Jacobi polynomials, convolution identity, 84
Jacobi triple product identity, 145, 152
Jacobi weight, 168, 179, 190
Jacobi-identity, 27
Jacobi-Trudi identity, 164
jump, 170
jump matrix, 172

k-Bessel function, 107
k-exponential kernel, 105
k-Laplacian, 100
`kfreerec`, 6
Kostka coefficients, 161
Kummer's theorem, 12

Laguerre polynomials, 7, 14, 17, 18, 158
large parameter, 214
largest zeros, 208
leaf node, 72
Legendre polynomials, 6, 10, 11, 13, 17, 18
lens, 185
Lie algebra, 27
Lie algebra $\mathfrak{su}(1,1)$, 31
Lie algebra $\mathfrak{su}(2)$, 31
Lie algebra homomorphism, 28
Lie algebra representation, 28
linear CMS model with harmonic confinement, 110
Liouville's theorem, 169
little q-Legendre polynomials, 19, 22
local hyperasymptotics, 217
lowest weight vector, 32

Macdonald identities, 146
Macdonald's identity, 100
Macmahon box theorem, 163
major index, 155
Maple, 1, 14, 22
matching condition, 197, 202
matrix Riemann-Hilbert problem, 172, 177
Mehler-Heine formula, 208
method of trees, 73
model, 40
modified Bessel functions, 168, 200
modified Jacobi weight, 168, 179, 196
moments, 156, 159
Morera's theorem, 169
Motzkin path, 157
multiplicity function, 97

non-commuting variables, 153
nth root asymptotics, 183

optimal number, 219
optimal truncation, 216
orientation anomaly, 224
oriented curves, 169
orthogonal Lie algebra, 35
orthogonal polynomials, 155, 172, 177
orthogonal polynomials on the simplex, 86
orthogonality relations for Clebsch-Gordan coefficients, 48
oscillatory behavior, 195
overlap coefficient, 72

p-lattice path, 156
parametrix, 168, 196
partition, 138
partition function, 143
Pascal triangle, 140
permutations, 158
Petkovšek's algorithm, 14
Pfaff-Saalschütz sum, 12, 38, 153
plane partitions, 162, 164

pochhammer, 3
Pochhammer symbol, 3, 33
positive discrete series representation $\mathcal{D}_k$ of $\mathfrak{su}(1,1)$, 52
positive maximum principle, 125
positive minimum principle, 118
positive operator, 117
positive subsystem, 96
projection, 45

q-binomial coefficient, 139
q-binomial theorem, 139, 141, 145, 153
q-Chu-Vandermonde theorem, 19
q-difference equations, 20
q-difference operator, 20
q-Dyson conjecture, 148
q-hypergeometric series, 19
q-Pochhammer symbol, 19
q-series, 138
q-Zeilberger's algorithm, 19
qsum, 19, 20
qsumdiffeq, 20
qsumrecursion, 19
quotient module, 30

Racah coefficient of $\mathfrak{su}(1,1)$, 68
Racah coefficients, 58
Racah polynomial, recurrence relation, 76
Racah polynomials, 10, 64
Ramanujan's congruences, 143
random matrix theory, 168
rank of a partition, 145
rational certificate, 13
re-expansion, 219, 225, 232
realization, 40
rec2ortho, 21
rechyper, 14
recurrence relation for dual Hahn polynomials, 52
recurrence relation for Racah polynomials, 76
reflection, 95
reflection group, 95
Regge array, 47, 56, 60
regular graph, 79
removable singularity, 169, 191, 204
representation, 28, 161
representation D_j, 32
representation of sl_2, 141
resurgence, 220, 225
REtoDE, 21
retode, 21
REtodiscreteDE, 21
REtoqDE, 22
Riemann-Hilbert problem, 168, 170, 171
Riemann-Lebesgue lemma, 110
Rodrigues formula, 17
rodriguesdiffeq, 17
rodriguesrecursion, 17
Rogers-Ramanujan identities, 148, 151
root system, 95, 146
rotation graph of binary coupling trees, 78
rotation on binary coupling trees, 78

Schur function, 161
semigroup, Feller-Markov, 124
semigroup, strongly continuous, 121
set partitions, 158
simple $\mathfrak{g}$-module, 29
SimpleRE, 5
simplex, 85
$\mathfrak{sl}(2,\mathbb{C})$, 27
$\mathfrak{sl}(n,\mathbb{C})$, 27
smoothing of the Stokes phenomenon, 217
$\mathfrak{so}(2n)$, 35
$\mathfrak{so}(2n+1)$, 35
$\mathfrak{sp}(2n)$, 34
steepest descent method, 168, 181
Stirling numbers, 158
Stokes line, 216, 229
Stokes multiplier, 230, 236, 238–240
Stokes phenomenon, 216, 223, 225
Stokes phenomenon, smoothing of, 217
strong asymptotics, 183, 193
strongly continuous semigroup, 121
structure constants, 28
$\mathfrak{su}(1,1)$, 31, 52
$\mathfrak{su}(2)$, 31
subalgebra, 27
sumdiffeq, 13
sumrecursion, 9
Sumtohyper, 5
sumtools, 3, 5, 8
symmetric chain, 142
symmetric functions, 160

symplectic Lie algebra, 34
Szegő condition, 183, 190
Szegő function, 183, 189

tensor product of $\mathfrak{g}$-module, 29
terminating hypergeometric series, 36
three-term recurrence relation, 156, 177
top node, 72
triad, 45

uncoupled basis, 41, 48, 57
unimodal, 140, 163

van der Corput-Schaake inequality, 104
Vandermonde determinant, 161
Vandermonde's summation formula, 37
Verma module, 29

V_λ, 33

weak convergence of zeros, 183
weight, 156
weight vector, 32
Weyl chambers, 96
Weyl denominator formula, 162
Weyl group, 147
Whipple's ${}_3F_2$ transformation, 38
Whipple's transformation, 10
Wilson polynomial, 67
Wilson polynomial, convolution identity, 82
WZ-method, 13
`WZcertificate`, 13

Zeilberger's algorithm, 9, 14